ROBUST SYSTEMS THEORY APPLICATIONS TO MACROECONOMIC STABILIZATION PROBLEMS

A Thesis in
Electrical Engineering
by
Muhittin Yilmaz

© 2005 Muhittin Yilmaz

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

December 2005
The thesis of Muhittin Yilmaz was reviewed and approved* by the following:

Mario Sznaier  
Professor of Electrical Engineering  
Thesis Advisor, Chair of Committee

Constantino Lagoa  
Associate Professor of Electrical Engineering

Octavia Camps  
Associate Professor of Electrical Engineering

Abdullah Yavas  
Professor of Business Administration

W. Kenneth Jenkins  
Professor of Electrical Engineering  
Head of the Department of Electrical Engineering

*Signatures are on file in the Graduate School.
ABSTRACT

This research involves the use of robust systems theory tools to deal with identification, model uncertainty and optimization in macroeconomic stabilization problems, in order to avoid Lucas Critique.

This is accomplished by using time domain based robust semi-blind identification to obtain models of macroeconomic systems of interest, as well as a bound on the identification error. By using past inputs to reflect the effect of the system initial conditions and convex relaxations, the original nonconvex model identification and (in)validation problems are reduced to jointly convex problems.

Linear Fractional Transformations (LFT) are used to optimally design a robust controller when model uncertainties and performance objectives can be treated in structural uncertainty framework. This LFT structure enables to achieve robust performance under simultaneous model uncertainty and performance objectives requirements.

These results are extended to time varying models by developing an LPV framework for semi-blind model identification and validation. The potential of this approach is illustrated by designing a robust controller for a macroeconomic stabilization problem, i.e., inflation targeting problems. Since the LPV framework can incorporate the macroeconomic dynamics changes for different equilibrium points and schedule controllers accordingly by utilizing linearly varying parameters, it answers the famous Lucas critique.
TABLE OF CONTENTS

LIST OF TABLES vi
LIST OF FIGURES vii
LIST OF SYMBOLS ix
ACKNOWLEDGMENTS xi

CHAPTER 1: Introduction 1
  1.1 Background ........................................... 1
  1.2 Lucas Critique ........................................ 7
  1.3 Problem Statement ...................................... 7
  1.4 Organization of the thesis ................................ 9

CHAPTER 2: Semi-blind Model Identification 11
  2.1 The Rudebusch and Svensson Model ...................... 11
  2.2 Robust Identification .................................... 13
    2.2.1 Semi-blind $\ell_1$–Identification ................... 18
  2.3 Macroeconomic Semi-blind Modeling ...................... 21
  2.4 Worst-case Prediction Error Bounds ...................... 24
  2.5 Conclusions and Future Research ......................... 26

CHAPTER 3: Semi-blind Model (In)validation 27
  3.1 Model (In)validation ...................................... 27
  3.2 Semi-blind Model (In)validation ......................... 29
  3.3 A Deterministic Convex Relaxation ...................... 32
  3.4 Macroeconomic Semi-blind Model Invalidation ........... 35
  3.5 Conclusions and Future Research ......................... 36
LIST OF TABLES

2.1 Rudebusch and Svensson’s (RS) two equation model variables . . . 12
3.1 Model Invalidation Results for the same prediction error bound . 36
3.2 Model Invalidation Results for the same uncertainty bound . . . . 37
LIST OF FIGURES

1.1 The Linear Parameter Varying control framework. .................. 9
2.1 The RS model simulations(inflation and output gap as subplots). 13
2.2 The block diagram for the $\ell_1$-identification with initial states. 18
2.3 The block diagram for the semi-blind $\ell_1$-identification. ........ 20
2.4 The Semi-blind $\ell_1$-Identification Simulation. ....................... 23
2.5 The RS model and the $\ell_1$-Identification forecasting errors. .... 24
3.1 The model invalidation setup. ........................................ 28
3.2 The semi-blind model invalidation setup. ............................. 30
3.3 The semi-blind model invalidation convex relaxation setup. ....... 32
4.1 The basic LFT building blocks. ....................................... 39
4.2 The basic LFT framework. ............................................ 39
4.3 The LFT representation of an uncertain parameter. .................. 41
4.4 The Robust Control framework. ...................................... 43
4.5 The $D$-scaling. ...................................................... 46
4.6 The control system interconnection structure. ....................... 49
4.7 The uncertainty($W_G(s)$) and performance($W_e(s)$) weighting func-
tions. ................................................................. 50
4.8 The macroeconomic stabilization problem in robust control frame-
work. ......................................................................... 51
4.9 The $\mathbb{H}_\infty$ and $\mu$ bounds. .................................... 51
4.10 The macroeconomic stabilization closed loop system. ............... 52
4.11 The closed loop system simulations. .................................... 52
4.12 The optimal inflation tracking and FED fund rate changes. ....... 53
5.1 The semi-blind LPV Identification Framework. ....................... 56
5.2 The LPV Framework Identification Data. ............................... 59
5.3 The historical data and the LPV identification-forecasting inflation outputs. ................................................................. 61
5.4 The semi-blind LPV Model (In)validation setup. ...................... 62
5.5 The basic LPV tracking problem block diagram. ...................... 67
5.6 The Linear Parameter Varying Synthesis Framework. ............... 68
5.7 Macroeconomic stabilization problem setup. ............................ 70
5.8 The LPV inflation targeting system outputs. ........................... 72

D.1 The LFT Small-Gain Theorem Framework. .............................. 88
LIST OF SYMBOLS

\( \sigma(A) \)  \hspace{1cm} \text{Singular value of a matrix } A. \\
\( \bar{\sigma}(A) \)  \hspace{1cm} \text{Maximum singular value of a matrix } A. \\
\( A^T \)  \hspace{1cm} \text{Conjugate transpose of a matrix } A. \\
A > (\geq)0 \hspace{1cm} A = A^T \text{ is positive (semi)definite matrix, i.e., } x^T A x > (\geq)0 \quad \forall x \in \mathbb{R}^n, x \neq 0. \\
\emptyset \hspace{1cm} \text{Empty set.} \\
\cap \hspace{1cm} \text{Intersection.} \\
\cup \hspace{1cm} \text{Union.} \\
| \hspace{1cm} \text{Identity matrix.}
The Banach space of vector-valued real sequences equipped with the $p$-norm, i.e., $||x||_p = (\sum_{i=0}^{\infty} ||x_i||^p_p)^{\frac{1}{p}}$ for $p \in [1, \infty)$ and $||x||_\infty = sup_i ||x_i||_\infty$.

The Lebesgue space of complex-valued matrix functions essentially bounded on the unit circle, equipped with the norm $||G||_\infty = sup_{|z|=1} \sigma(G(z))$.

The subspace of functions in $\mathcal{L}_\infty$ with bounded analytic continuation inside the unit disk, equipped with the norm $||G||_{\infty, \rho} = sup_{|z|<\rho} \sigma(G(z))$.

The space of transfer matrices in $\mathcal{H}_\infty$ equipped with the norm $||G||_{\infty, \rho} = sup_{|z|<\rho} \sigma(G(z))$.

The open unit ball in normed space $\mathcal{X}$.

The open $\gamma$-ball in $\mathcal{X}$.

Toeplitz operator, i.e., $T_G : l^\infty[0, \infty) \rightarrow l^\infty[0, \infty)$, associated with an $l^\infty$ stable system $G$.

Hankel operator, i.e., $\Gamma_G : l^\infty(-\infty, -1] \rightarrow l^\infty[0, \infty)$, associated with an $l^\infty$ stable system $G$. 
ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Mario Sznaier, for his encouragement and endless ideas and, the committee members, Dr. Yavas, Dr. Lagoa and Dr. Camps for their service, valuable feedback and support throughout this research. I am especially indebted to Dr. Yavas for his support and forward-looking view of the research on macroeconomic concepts.

I also would like to express my thanks to Wenjing Ma for her support during the programming of the semi-blind identification and model validation topics.
DEDICATION

This research is dedicated to my parents, my relatives and especially my father who passed away during the research. Without their support, encouragement and good wishes, it would have been almost impossible to finish the study.

This study is also dedicated to all open-minded researchers who stand against, the way I call it, scientific bigotry. Recalling the hardship I faced at the beginning of this study, I believe that all open-minded researchers deserve huge acknowledgement and appreciation for their openness and willingness to venture their main course of research for the sake of exploring new horizons.

Last but not the least, this study is dedicated to my beloved wife who spent a lot of effort, showed a lot of courage and support, to make this research accomplish its goals.
Chapter 1

Introduction

1.1 Background

When Milton Friedman[1] recommended that policy makers acknowledge model uncertainty, he actually meant that the subjective or objective probability distributions can not characterize the type of uncertainty. He expressed that he knew too little about the economic dynamics and did not advocate using prior beliefs about different specifications of dynamic monetary models[3]. Frank Knight [2] defined risk and uncertainty differently by emphasizing that risk meant outcome for known probabilities and uncertainty meant outcome for unknown probabilities. Brainard[4] suggested that whenever there is uncertainty about the true parameters of a model, policy makers should be cautious, and move monetary policy instruments by less than what they would in the absence of uncertainty.

Until recently, monetary policy rules have been justified mainly by using a Bayesian approach. In this approach, a prior probability distribution is assigned to each model parameter and a monetary policy is obtained by minimizing the expected loss by using these probability distributions. However, assigning a prior probability distribution to each model parameter implies that the designer actually knows the model or has very good idea of the parameters. However, practical macroeconomic models have many parameters and assigning priors to each of them is not easy and practical. On the other hand, using simple and uninformative priors, i.e., flat distribution, just to overcome this problem is unjustified. Moreover, due to its parametric nature, the Bayesian approach can not handle situations
where there could be different models for different states of the economy. Now, since one has to deal with the model as a whole, it is not possible to assign priors for parameters. Assigning priors to show varying parameters\cite{5} and incorporating learning to update posteriors\cite{6} are more advanced treatments of uncertainty in Bayesian approach. The Bayesian approach suggests that model uncertainty results in a less aggressive monetary policy than the one without model uncertainty. Peter Huber stated in his book Robust Statistics\cite{7},

... one justifies their use by appealing to a vague continuity or stability principle: a minor error in the mathematical model should cause only a small error in the final conclusions. Unfortunately, this does not always hold. During the past decade one has become increasingly aware that some of the most common statistical procedures are excessively sensitive to seemingly minor deviations from the assumptions ...

— Peter Huber

and stated that assumptions that are not exactly true are mathematically convenient ways to represent fuzzy knowledge.

Although it was once popular, optimal control theory suffered from the fact that its robustness properties were not satisfactory, yielding instability in some cases, when there is even a small uncertainty. To overcome these deficiencies, modern robust control theory was born to ensure system robustness properties against model uncertainty and disturbances. Starting with Zames\cite{8} in early 80s, precise definitions for uncertainty, performance and important developments\cite{9, 10} yielded powerful and mathematically tractable techniques such as $H_\infty$\footnote{1$H_\infty$ approach is a well-established worst-case control design theory and it tries to minimize the ratio of $l_2$-bounded output over $l_2$-bounded input, which is equal to the induced $\infty$-norm of the transfer function between the mentioned input-output, under some stabilization constraints.} framework, $\mu$-controller synthesis\footnote{2This is an improvement, yet harder to solve, over $H_\infty$, in which uncertainty is assumed to be structured.} and $\ell_1$-optimal control. While ”robust” means the guaranteed optimal performance and the system stability under uncertainty in robust control theory, the insensitivity of the system response to minor changes in parameters is called ”robust” in statistics. In spite of its statistical basis, Risk Sensitive control\cite{11} studied a standard Gaussian setup with no uncertainty and obtained
the same formulae in some cases as the ones obtained by $H_\infty$. Searching for a robust policy rule is not new but using robust control theories to achieve robustness under different equilibrium points is a pretty new approach.

A number of researchers have studied the relationship between model uncertainty and the degree of monetary policy attenuation. In addition, several researchers have used robust control to deal with model and shock uncertainty. McCallum[12] studied robustness for misspecified models by applying the same rule for different models and analyzing the outcomes. He emphasized on the disagreement of a common structural macroeconomic model and tried to find a policy rule that works well over plausible models, instead of finding an optimal policy rule for any particular model. Thomas Sargent proposed using unstructured uncertainty to deal with misspecified models from a robust control perspective[13]. Hansen and Sargent[14] used robust control to study Ramsey problems in different setups.

Aaron Tornell[15] used an $H_\infty$ approach in forecasting asset price anomalies and stressed that uncertainty is modeled by stochastic processes with known or can-be-learned properties and that agents maximize an expected utility in a Rational Expectations(RE) framework. Here, uncertainty is modelled as an unknown sequence with bounded $l_2$ norm, and the ratio of discounted utility to the $l_2$-norm of disturbances is used, in an $H_\infty$ context, to obtain closed-form decision policies. Following his claim that existing $H_\infty$ control, which should consider the interdependency of prices and agents’ decision in equilibrium, is not applicable in economic problems, he constructed a competitive equilibrium of an exchange economy. He also obtained better stock price forecasts by using dividend data from 1871-1996 in $H_\infty$ approach than Rational Expectations forecasts due to $H_\infty$ sensitivity to news. He stressed the main difference: $H_\infty$ policies work well under norm-bounded uncertainties whereas Rational expectations produce the best performance when no uncertainty exists. He found that although RE and $H_\infty$ setups are different, their forecasting formulas are quite similar. He stated that since $H_\infty$ was developed for engineering problems, there are differences in engineering and economics problems such as the different performance indexes ($H_\infty$ includes mean square of the state while in economics payoffs are functions of noisy observations) and equilibria characteristics and formulas should be adjusted.
Onatski and Stock argue that the minimax approach of robust control provides a general and tractable alternative to the conventional Bayesian decision theoretic approach and study a two-equation US macroeconomic model by using robust control tools[16]. They also used structured model uncertainty under different types of problems($H_{\infty}$, $\ell_1$, etc.) to analyze the policy rule characteristics. They explicitly assume the Taylor rule\(^3\) as the feedback rule and compared the shock and model uncertainty(both structured and unstructured) cases with the typical Taylor policy rule in terms of aggressiveness of the policy in case of uncertainty. Their main results are that the robust rules are more aggressive than LQG rules and the robust rules vary depending on the uncertainty formulation. Tetlow and Von zur Muehlen [17] studied both a structured and unstructured model and shock uncertainties for forward-looking models both in time and frequency domains(Forward-looking models involve future values of states, i.e., estimation of future variables for current state calculations). Another difference is that economists tend to separate model and shock uncertainty to better utilize uncertainty structure for the economy. In general, unstructured uncertainty contains both model and shock uncertainty. The certainty equivalence principle says that the optimal solution is independent of noise(shock) coefficient. The certainty equivalence principle needs a quadratic objective function, linear transition and shocks independent of states. Thus, although the certainty principle holds for shock uncertainty case where shocks are independent of states, it does not hold for model uncertainty case, i.e., misspecified models. Giannoni [30], on the other hand, studied parameter uncertainty and robust optimal rules based on zero-sum two player games. His work didn’t involve any robust control tools.

Inflation forecast and policy choices are very important questions in terms of theoretical and empirical aspects. US Federal Reserve bank(FED) have chosen policies toward having a low and stable inflation, i.e., inflation targeting, since 80s in which the future inflation levels need be forecast accurately. However, this is difficult due to complexities of inflation indicators. It is difficult, for instance, to determine the useful indicators(money supply, gold prices, etc.) to forecast future inflation.

\(^3\)The generalized Taylor rule is an estimation of policy rule over years and it is
\[ i_t = g_\pi \pi_t + g_y y_t, \]
where $i_t$ is the nominal interest rate(the monetary authority tool), $\pi_t$ is the average of last four quarter inflation rates, $y_t$ is the output gap, and constants $g_\pi = 1.5$ and $g_y = 0.5$.
inflation. It is also difficult to establish how a policy change affects the future inflation path since there is no exact system dynamic representation between the FED instrument and inflation. (i.e., what is the correct policy if inflation is forecast to be different than the target?) Stephen G. Cecchetti [18] studies the feasibility of inflation targeting and finds that it is hard both to forecast inflation due to structural breaks and to obtain a system dynamic and concludes that nominal income targeting reduced the changes in real income, while resulting only in a minor increase in inflation. He then considers to forecast inflation at a different future time by using different indicators due to the implications of economic theory. He finds that commercial forecasts are poor, that the forecast performance by Global Insight company[^4] declines as the horizon increases. He then moves on to find possible candidate indicators by using their correlation with inflation utilizing a sixth order lag equation. He uses a number of candidate inflation indicators, including several commodity price indices, the price of gold, monetary aggregates, the price of oil, interest rates and spreads, a wage index, unemployment, etc. Following the work of Niemira and Klein[19] and Webb and Rowe[20], he concludes that candidate variables are correlated with inflation for different horizons, that different sample period results are different, i.e., inflation dynamics is changing. This change is modeled as structural breaks subject to Lucas critique[31] by Cecchetti[18], while Caskey[21] studies this change by a linear model with Bayes learning. Another detailed analysis by Stock and Watson[22] included a decade long literature review and used 38 mainly asset price candidate indicators to forecast output and inflation. At the end, they came up with the same result: some asset prices are good to predict inflation or output for some periods for some countries but not all the times.

Economic theory analysis suffers when the available mathematical tools are not sophisticated to explain real-life paradigms. Since robust control tools were not available to handle model variations, optimal control applications were suddenly ceased due to Lucas Critique which argued the plausibility of optimization of varying systems by using fixed models and many stochastic methods were extensively used for economic analysis, etc. For example, rational expectations theory is very

[^4]: Global Insight company is located at http://www.globalinsight.com, contains historical time series data and does analysis and forecasting of future variables.
famous among economists. But, it astonishingly excludes model misspecification so that, when this theory was developed, model misspecification issues were ignored. As stated[33], misspecified models overwhelms the rational expectations equilibrium concept and forces it to reset its common model requirements. Since it is now known how to handle model uncertainty, rational expectations theory can be further explored.

The inevitable disagreement on macroeconomic models, variables and structures suggests that robust control, not necessarily having an articulately formulated economic process for the underlying dynamics, is a very viable compromising solution. By using robust control tools, the actual system can be modeled as a nominal system and a norm bounded uncertainty. Control oriented identification in terms of $\ell_1$ and $H_{\infty}$ are well established and very suitable tools for robust control optimization problems [23], [24], [25], [26], [27], [28]. Moreover, the uncertainty structure can be exploited to increase the efficiency of the optimization. Structured uncertainty optimization results are at least equal to, if not better than, the unstructured uncertainty optimization. Due to the very changing nature of the overall economy, one nominal model and a bounded uncertainty will yield a very conservative design. Since unstructured and structured uncertainty cases by using Taylor rule for backward and forward-looking statistical models have been studied[14, 16], LPV framework is the next step to better use the robust system theories in deterministic setup. In robust control linear parameter varying(LPV)[29] framework, changes in economic dynamics as well as adjustments in the corresponding feedback control law due to authorities policy change, individual’s preference change, etc., are achieved by using linearly varying parameters, assumed to be measured in real-time, to indicate different equilibrium points. Since the LPV framework efficiently utilizes convexity properties of the possible models, scheduling among different LPV controllers is ensured to be smooth transition. This says that the LPV controller synthesis is done once and can be used for all possible system models since the policy rule can adjust against all possible cases.
1.2 Lucas Critique

Robert E. Lucas, Jr.[31], who won the Nobel Prize in Economic Sciences in 1995, have contributed to macroeconomic theory since 1970s. One of his main work is known as Lucas Critique: although monetary transmission mechanisms, economic dynamics models, can be identified by using parameter estimates yielding a model with fixed parameters, Lucas argued, if the economic authorities are pursuing completely different policy, these parameters in the model should be different to accommodate the changing nature of the economic dynamics. Economists have built models with parameters, a set of equations, to represent the economy, expectations, firms, preferences and technology. Since these all depend on particular policy choices, the model parameters are likely to change systematically with the policy change. If the model doesn’t change, it implies that the model will be fixed for different policy regimes, to which Lucas responds as

... Everything we know about dynamic economic theory indicates that this presumption is unjustified. ...

— Robert Lucas

and uses the idea that models are based on decision rules of economic participants and models should reflect the changes in people’s future expectations due to the current policy changes. Lucas critique virtually ceased the work on the macroeconomic optimization by using optimal control with fixed system models, which was popular to use for policy evaluation until 1970s and yielded to Rational expectations and forward-looking models. However, there exist papers arguing against the theoretical and/or empirical foundations of the Lucas critique [36].

1.3 Problem Statement

Since engineers prepare for the worst case while economists usually average outcomes, it is argued that one of the differences between economists and engineers is that when there are 100 bridges and if one of them collapses, while engineers scream economists feel lucky. When dealing with misspecified economic transmission mechanisms, robust system theories seem very appealing to determine robust monetary policy rules. As some argue, it is a valuable alternative to Bayesian
analysis[17]. Noting this point, a robust control based deterministic optimization framework immune to Lucas critique can be proposed to find quantitative answers for macroeconomic stabilization problems qualitatively stated as follows;

**Problem 1.3.1 (Macroeconomic Stabilization Problems).** When there are undesired changes or exogenous shocks in the economy such as a jump in unemployment or inflation, a decrease in gross domestic product, etc., in order to correct the situation optimally in foreseeable future, what should financial authorities, namely the government and the central bank, do, when should they do, by how much should they do and how long should they do?^{5}

Macroeconomic stabilization problems can be quantified by using the LFT\(^6\) block representation in Fig.1.1, where \(y\) denotes the outputs to be used at the controller, \(u\) denotes the controller input to the system, financial authorities’ policy tools, namely short-term interest rate, government spending and tax rates, \(\omega\) denotes the exogenous inputs, \(z\) denotes the outputs to be used to measure performance, \(\phi\) denotes the linearly varying parameter(s) with bounded rate of change, assumed to be measurable in real-time, \(M(\phi)\) denotes the affine parameter-dependent system to accurately reflect the varying nature of economic transmission dynamics, \(K(\phi)\) is the LPV controller to be designed for given performance objectives and \(\Delta\) denotes the norm bounded model uncertainty.

The LPV framework, shown in Fig.1.1, can be utilized to mathematically present the proposed macroeconomic stabilization problems;

**Problem 1.3.2.** Find a parameter dependent controller \(K(\phi)\) to achieve the performance objective(s) \(Q = \|f(\omega,z)\|_q \leq \gamma, \text{ subject to } \|\Delta\|_p \leq \delta\) and affine parameter-dependent \(M(\phi)\) by using output and/or state feedback.

LPV framework will include an affine parameter-dependent macroeconomic dynamics identification with uncertainty, model validation and optimization by de-

---

^{5}Stephen G. Cecchetti [18] asks about inflation forecast and policy options and tries to answer this question in statistical framework by minimizing the mean-square error of some variables such as commodity prices, Fed funds rate, etc. by calculating the impulse responses of the corresponding shocks to the policy innovations over some horizon. He concludes that forecasting inflation is a difficult task over different time horizons and that the inflation-monetary policy instrument relationship is too sensitive to obtain accurate estimates.

^{6}LFT representation is a very powerful as well as flexible way to represent systems with uncertainty in a block diagram, where an augmented plant possibly including robustness and performance weights, a controller and uncertainty appear as separate blocks.
signing a controller. Since LPV framework can possess better system models via varying parameters for different operating points, the designed controller will not be as conservative as when the unstructured full block model uncertainty is assumed.

Since economists are using statistically established models (both time-series and lagged-values)\(^7\), deterministic approaches can be shown to be useful. Also, Economists try to understand how different types of uncertainty and problem setups will affect monetary authority’s behavior. So far, although there is no consensus, it is believed that robust rules are less aggressive than rules in the absence of uncertainty. But, Giannoni [32] points out that the model and the type of the uncertainty considered will mainly determine the aggressiveness of the policy rule.

Due to the size of economic activities, i.e., trillions of dollars, and the importance of the welfare of a society, improving the macroeconomic stabilization problem performance even by 1% will translate to saving significant amount of money or having lower rate of unemployment, etc.

### 1.4 Organization of the thesis

Chapter 1 contains historical background for control applications to macroeconomic system stabilizations involving uncertainty, Lucas critique and problem def-

---

\(^7\)Economists develop theories and find a statistical model by using variables in the theory. However, engineering perspective suggests that all important variables should be used as long as the problem is tractable.
In Chapter 2, a semi-blind\textsuperscript{8} $\ell_1$-identification technique is proposed. The time series data for the US economy is used to identify a model with a bounded uncertainty. This analysis shows that robust $\ell_1$-identification can successfully be used to forecast economic variables.

In Chapter 3, semi-blind model (in)validation for the macroeconomic model is presented. Since the macroeconomy is already progressing for the identification data range, input initial conditions are used to minimize the error and uncertainty bounds so that the model is validated for a longer data range. Simulations proved that large uncertainty bounds are partially coming from the system error caused by the initial conditions.

In Chapter 4, LFT representation and robust controller design for structured uncertainty case are explained and applied to an economic model to interpret the attitude of the Federal Reserve Bank (FED) against uncertainty. The results show that FED reacts less aggressively than when there is no uncertainty.

In Chapter 5, a Linear Parameter Varying (LPV) framework for semi-blind affine parameter-dependent system identification with uncertainty, model validation and optimal controller design is introduced and applied to a monetary stabilization problem, i.e., inflation targeting problems. Since this framework recognizes model variations and adjusts both model and the corresponding feedback law, this synthesis is the answer to Lucas critique.

In Chapter 6, overall conclusions and directions for further research are given.

\textsuperscript{8}Semi-blind refers to using unknown initial conditions.
Semi-blind Model Identification

A semi-blind ℓ₁-identification framework is introduced and applied to find a model of a subsystem of the US economy.

2.1 The Rudebusch and Svensson Model

The Rudebusch and Svensson (RS) two-equation empirical model of the US economy \[39\] are given in Eqns.2.1 and 2.2 and related variables are explained in Table 2.1.

\[
\begin{align*}
\pi_{t+1} &= 0.7\pi_t - 0.1\pi_{t-1} + 0.28\pi_{t-2} + 0.12\pi_{t-3} + 0.14y_t + \epsilon_{t+1} \quad (2.1) \\
y_{t+1} &= 1.16y_t - 0.25y_{t-1} - 0.1(i_t - \pi_t) + \epsilon_{y_{t+1}} \quad (2.2)
\end{align*}
\]

Eqn.2.1 represents the Phillips curve, i.e., an aggregate supply equation, the relationship between the inflation and the output gap. Adaptive representation of inflation expectations are incorporated by using the lags of inflation. Eqn.2.2, IS curve or an aggregate demand equation, relates the output gap to its own lags and to the ex-post real interest rate. The RS model seems to have good properties; the model monetary policy actions affect output first and then inflation, it is linear and tractable, its structure\(^1\) is typical of central bank models. Its practical policy-oriented aspect appeals many central bankers, who conduct actual

\(^1\)Using short-term interest rate as the policy instrument, defining the model in terms of output gaps instead of output growth rates and including a Phillips curve with autoregressive expectations consistent with the natural rate hypothesis
Table 2.1. Rudebusch and Svensson’s(RS) two equation model variables

\[ \pi_t = 400. [\ln(p_t) - \ln(p_{t-1})] \]  
the quarterly inflation in the Gross-Domestic Product chain-weighted price index\((p_t)\) in percent at an annual rate,

\[ \bar{\pi}_t = 0.25 \sum_{k=0}^{3} (\pi_{t-k}) \]  
the last 4-quarter average of the inflation,

\[ y_t = 100 \left( \frac{q_t - q^*_t}{q^*_t} \right) \]  
the percentage gap between actual real\((q_t)\) and potential GDP\((q^*_t)\),

\[ i_t \]  
the quarterly Fed funds rate in percent at an annual rate

\[ \bar{i}_t = 0.25 \sum_{k=0}^{3} (i_{t-k}) \]  
the 4-quarter average of the Fed funds rate and i.i.d zero mean disturbances.

monetary policies. Although it is backward-looking and subject to Lucas critique, it is used by many central bankers\[46\]. The RS model equations were estimated individually by ordinary least squares for the quarterly data between 1961.1-1996.2 period for the US economy. During the original model estimation, model variables were demeaned so that constants do not appear in the equations. Rudebusch and Svensson were able to justify their model performance results with respect to MPS model\[48\] and with respect to historical simulations\(^2\) given on Fig.2.1.

Rudebusch and Svensson tested the performance of policy rules in an inflation targeting monetary policy regime. They consider two classes of rules: instrument and targeting rules. The monetary policy rule is stated as a function of available information in the instrument rule class. One famous example of this class is Taylor’s instrument rule\(^3\), i.e., the nominal interest rate is given in terms of current inflation and output\[44\]. Not being used directly, instrument rules are used to provide baseline to the actual policy. On the other hand, targeting rules are used by assigning a loss function over deviations of an intermediate or a goal variable from a target level. They studied different instrument and targeting rules and presented their results on output and inflation changes under various policy rules. They conclude that some simple targeting and instrument rules using inflation

\(^2\)Output gap simulations, standard errors and other statistics are omitted here due to the scope of this study.

\(^3\)Taylor rule is briefly explained on page 4.
forecast behave better to minimize the loss function than the optimal rule while inflation targeting policy rules behave worse than expected. From the control theory point of view, targeting rules are relevant to finding feedback laws in an optimized framework and instrument rules are relevant to determining the system output and the system input. It should be noted that a targeting rule becomes an implicit instrument rule for a particular model after the system optimization.

As seen from the Fig.2.1, RS model fits the data well. However, the RS model requires a large number of data points while not utilizing the \textit{a-priori} information about the system under consideration. Moreover, any appropriate conclusion can not be deemed about modeling and future data forecasting. This dissertation partly aims to answer these pitfalls by pursuing a deterministic, set membership approach to model identification.

### 2.2 Robust Identification

System identification associated with a process is a basic building block for analysis and synthesis purposes where finite, partial and corrupted data is used to obtain a mathematical description of the corresponding system. Until the appearance of
systematic algorithms in late 80s, some ad hoc procedures were used\(^4\). The traditional identification approaches\([42, 43]\) involve statistical tools, a prescribed fixed model structure whose output is only corrupted by noise with known statistical properties and associated identification algorithms that use least squares methods to estimate the correct values of the parameters proposed in the presumably correct model structure. Moreover, since the stochastic noise is the only source for uncertainty while the prescribed model structure is taken as the correct description of the system in these algorithms, they are not appropriate to characterize mathematical models subject to model uncertainty. This is partly due to the observation that the stochastic noise description may not be the only source of uncertainty, perhaps, the assumption about the model structure may be unrealistic and the model structure-parameters may be changing at different operating conditions.

Control-oriented identification techniques assume deterministic, bounded yet unknown noise properties and aim to provide a nominal model with a quantifiable worst-case modeling error. Therefore, deterministic norm bounded identification procedures were proposed by using the set membership formulation\(^5\) to initiate the robust control design. The emerging field is called as robust identification and was first proposed by Helmicki\[24\]. The objective in robust identification methods is both to identify a linear time-invariant (LTI) nominal model convergent to true model and to determine the worst-case identification error. Robust identification techniques are built on worst-case behavior of the process over sets of candidate models \(S\) and noise \(N\), two a-priori assumptions, and actual system time series \((u,y)\), a-posteriori data.

Information consistency establishes the plausibility of the a-priori assumptions about the set of candidate models \(S\) and the noise set \(N\) to a-posteriori data and considers the set of all possible models that could produce the a-posteriori data, in accordance with the class of systems \(S\) and the measurement noise \(\eta \in N\) such that

\[
\mathcal{T}(y) = \{g \in S | y_k = (g * u)_k + \eta_k\} \tag{2.3}
\]

\(^4\)For example, when a system modeling is too complicated from physical considerations, D\(_2\)O plant controller is designed by using a general model \[45\].

\(^5\)During '90s, the set membership approach received a lot of attention and a survey for system identification by using different set membership formulations can be found \[47\].
for some sequence $\eta_k \in \mathcal{N}, k = 0, 1, ..., N - 1$. The set $T(y) \subset \mathcal{S}$ is the smallest set of models indistinguishable with the available information.

**Definition 2.2.1 (Information Consistency).** The a-priori assumption $(\mathcal{S}, \mathcal{N})$ is consistent with the a-posteriori information $y$ if and only if the set $T(y)$ is nonempty\(^6\).

After consistency is established between a-priori assumptions and a-posteriori data, a nominal model and a model error computations can be done. Interpolation algorithms\(^{[41]}\) always give an identified model member of the consistency set and are guaranteed to converge to the true system model as the information is completed. Furthermore, by using the results on page 84 in Appendix B, it can be shown that these algorithms are optimal within a factor of 2, i.e., the worst-case identification error is, at most, equal to two times of the error calculated by hypothetical optimal algorithm.

Nonparametric identification may yield conservative results if a part of the system represents a clear structured behavior. In these cases, a parametric portion should be added to the identification procedure to exploit the a priori system knowledge. The a-priori assumptions for the two parameters $(K, \rho)$ of the candidate model set $\mathcal{S}$ represent the magnitude and smoothness of the candidate set of models. If a system has a peak in the frequency response, it should be taken care of by selecting a high $K$ value, which unfortunately may be very conservative. However, a low order model can be used to characterize this frequency peak and much smaller a-priori $K$ value can be used, yielding a much better result in terms of identification error. With regard to the macroeconomic problems, when identifying the system using FED funds rate as input and the inflation as output, there is no clear advantage in using a parametric component. However, if the difference of the FED funds rate is used as the input, this immediately implies that there is an integrator in the system. This integrator can be treated as the parametric portion of the system and whatever is not explained by this parametric part can be treated as the nonparametric part of the system. When the parametric

\(^6\)Obviously, $T(y) = \emptyset$ shows that any candidate model in $\mathcal{S}$ with noise in $\mathcal{N}$ cannot generate the measured a-posteriori data. Since the a-priori assumption is a quantification of engineering common sense or more popularly "a leap of faith", it does not guarantee that they will be coherent with a-posteriori data.
and nonparametric models are used during identification, it is referred as mixed parametric/nonparametric identification[53].

If the identification data is based on frequency domain experimental outputs, it is called $H_\infty$-identification. On the other hand, if the identification data is based on time series outputs, it is called $\ell_1$-identification. Both identification algorithms are linear\footnote{Any identification algorithm is linear as long as the identified model is a linear function of experimental outputs.} and depend on polynomial interpolation techniques.

$\ell_1$–Identification uses noisy time domain data points to generate a nominal system model and a worst case error. The corresponding identification algorithms can be obtained by utilizing time-domain interpolation methods[40]. The fundamental result for $\ell_1$ time domain identification comes from Carathéodory-Fejér interpolation problem and the main result is given in Lemma 2.2.1.

**Lemma 2.2.1 (Carathéodory-Fejér).** Given a matrix valued sequence $\{L_i\}_{i=0}^{n-1}$, there exists a causal, discrete-time, LTI operator $L(z) \in BH_\infty$ such that

$$L(z) = L_0 + L_1 z + L_2 z^2 + \ldots + L_{n-1} z^{n-1} + \ldots$$

if and only if $(T_L^n)^T T_L^n \leq I$.

**Proof:** See [51].

In robust identification, any operator $G$ is represented by a rational complex-valued transfer function

$$G(z) = \sum_{k=0}^{\infty} g_k z^{-k}$$

or by a minimal state space realization

$$G \equiv \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

where $G = C(zI - A)^{-1}B + D$ or $g_0 = D$, $g_k = C.A^{k-1}.B$ for $k \geq 1$. Moreover, Hankel, $H_G : \ell^\infty(-\infty, -1] \rightarrow \ell^\infty[0, \infty)$, and Toeplitz, $T_G : \ell^\infty[0, \infty) \rightarrow \ell^\infty[0, \infty)$, operators can be defined to show the past and current input mappings to the output, respectively, for each linear, discrete, time-invariant system $G$ and given
in Eqn.2.7. Finally, $T_G^N$ and $H_G^N$ denote the $(N \times N)$ size upper left finite Toeplitz and Hankel matrices, respectively.

\[
T_G = \begin{bmatrix}
g_0 & 0 & 0 & 0 & 0 & \ldots \\
g_1 & g_0 & 0 & 0 & 0 & \ldots \\
g_2 & g_1 & g_0 & 0 & 0 & \ldots \\
g_3 & g_2 & g_1 & g_0 & 0 & \ldots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\
\end{bmatrix} \quad H_G = \begin{bmatrix}
g_1 & g_2 & g_3 & g_4 & g_5 & \ldots \\
g_2 & g_3 & g_4 & g_5 & g_6 & \ldots \\
g_3 & g_4 & g_5 & g_6 & g_7 & \ldots \\
g_4 & g_5 & g_6 & g_7 & \ldots & \ldots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\
\end{bmatrix}
\] (2.7)

Related to the interpolatory algorithms, consider systems of the form:

\[
S \doteq \{ G(z) = H(z) + P(z) \}
\] (2.8)

where $H(z) \in BH_{\infty, \rho}(K)$ for some $\rho \geq 1$ and $P(z)$ represent the nonparametric and parametric components of the operator, respectively. Since $P(z)$ is the parametric part, it is assumed to belong to the class $\mathcal{P}$ of affine operators:

\[
\mathcal{P} \doteq \{ P(z) = p^T G_p(z), \quad p \in \mathbb{R}^{N_p} \},
\] (2.9)

where the $N_p$ components $G_p_i(z)$ of vector $G_p(z)$ are known, linearly independent, rational transfer functions. The following result gives a necessary and sufficient condition for two finite vector sequences to be related by an operator in the $S$ family.

**Lemma 2.2.2 (Existence of an operator).** [52] Given $K$, $\rho$ and two vector sequences $(u, y)$, there exists an operator $S \in S$ such that $y = S \ast u$ if and only if there exists a vector $h$ satisfying:

\[
M(h) = \begin{bmatrix}
K R^{-2} & (T^N_R)^T \\
T^N_h & K R^2 \\
\end{bmatrix} \succeq 0
\]

\[
y = T^N_u P p + T^N_u h
\] (2.10)

---

8 Since $H(z)$ corresponds to the approximation error during the finite expansion of the system by basis functions, this notation implies that the upper bounds for the magnitude and the time constants of the approximation error are available.

9 This could be interpreted as the identified system expansion $p$ in terms of predetermined basis functions $G_p$, which are expected to be linearly independent.
where \((P)_{k} = [g_{1}^{k} \ g_{2}^{k} \ \cdots \ g_{N_p}^{k}]\), where \(g_{i}^{k}\) denotes the \(k\)th Markov parameter of the \(i\)-th transfer function \(G_{pi}(z)\), \(h_{k}\) is the \(k\)th Markov parameter of the non-parametric component \(H(z)\), \(T_{u}^{N}\) is the lower Toeplitz matrix associated with the sequence \(u\), and \(R\) is given in Eqn.2.11.

\[
R = \begin{bmatrix}
1 & \rho & \rho^2 & \cdots & \rho^{N-1}
\end{bmatrix}.
\] (2.11)

In this case, all such operators \(S\) can be parameterized in terms of a free parameter \(Q(z) \in \mathcal{BH}_{\infty}\). In particular, the choice \(Q(z) = 0\) leads to the “central” model \(S_{\text{central}}(z) = H_{o}(z) + P^{T}G_{p}(z)\) where an explicit state–space realization of \(H_{o}(z)\) can be found for instance in [53].

### 2.2.1 Semi-blind \(\ell_1\)-Identification

Until now, the robust \(\ell_1\)-identification is done simply by ignoring the system initial conditions. However, if the system is already in progress, ignoring the initial conditions may cause large identification errors. With regard to macroeconomic system identification, effect of the FED funds rate prior to the time of interest should be encapsulated into some variables during the identification so that high identification errors may be avoided. Thus, a new identification problem with initial conditions can be proposed by using the setup in Fig.2.2.

![Figure 2.2](image)

**Figure 2.2.** The block diagram for the \(\ell_1\)-identification with initial states.
Problem 2.2.1 (Semi-blind \( \ell_1 \) identification). Given an unknown plant, a-priori sets of candidate models and noise \((\mathcal{S}, \mathcal{N})\) and a finite set of samples of the input \( u \) to the plant and its corresponding output \( y \) corrupted by additive measurement noise \( \eta \), find a model \( g \) compatible with both the a-priori information and the a-posteriori experimental data, that is \( g \in \mathcal{T}(y) \), where

\[
\mathcal{T}(y) \doteq \{ g \in \mathcal{S} : y_k - \sum_{i=0}^{k} g(i)u^+(k-i) - C_G (A_G)^{k-1} x_o \in \mathcal{N}, \ (k = 0, \ldots, N-1) \},
\]

for some \( x_o \) (2.12)

where

\[
G = \begin{pmatrix} A_G & B_G \\ C_G & D_G \end{pmatrix}, \ g_o = D_G, \ g_i = C_G (A_G)^{i-1} B_G
\]

By using the Fig.2.2,

\[
T^N_\omega = T^N_y - T^N_\eta
\]

can be written\(^\text{10}\). Substituting

\[
T^N_\omega = T^N_G T^N_{u^+} + T_{x_o} x_o\]

in Eqn.2.13, and by invoking Lemma 2.2.2, the Problem 2.2.1 reduces to a Bi-Affine Matrix Inequality (BMI) in terms of \( \{g_i\} \) and \( x_o \), both of which are unknown. However, BMIs generically lead to non-convex, NP-hard optimization problems. To avoid this difficulty, we propose a convex relaxation of the Problem 2.2.1 where the effects of unknown initial states \( x_o \) are replaced with the effect of unknown past inputs \( u^- \) which are assumed to belong to some set \( \mathcal{U} \). Since the model is a controllable-observable linear time-invariant system, there is a linear relationship between the system initial states and the system past input(s) such that optimization results should be similar whether the system initial states or the system past

\(^\text{10}\)Note that while \( y \) values are given, \( \omega \) and \( \eta \) are variables during identification.

\(^\text{11}\)Since the FED funds rate is the input variable, this assumption holds for the economic subsystem under consideration here. Moreover, the way the input is set up, i.e., the FED funds rate difference, is a perfect fit for this assumption.
inputs are used, provided that they are mapped to the system output properly. Consequently, the effect of the unknown initial states $x_o$ can be replaced by the effect of an unknown signal $u^- \in U$ acting in $(-\infty, -1]$, as shown in Fig.2.3.

\[ S = H(z) + P(z) \]

Figure 2.3. The block diagram for the semi-blind $\ell_1$-identification.

This substitution of variables leads to the following reformulation of equation (2.12):

\[ T(y) \doteq \{ g \in S : y_k - (H_g u^-)_k - (T_g u^+)_k \in \mathcal{N} \quad (k = 0, \ldots, N - 1) \}, \]

for some $u^- \in U$ \hfill (2.15)

where $u^+ \doteq \{u_0, u_1, \ldots, u_{N-1}\}$ and where $H_g$ and $T_g$ represent the Hankel and Toeplitz operators associated with the system $g$, respectively. $(H_g u^-)_k$ can be replaced by a new variable $x_k$ subject to the constraint that $x_k = (H_g u^-)_k$ for some $u^- \in U$. Assuming that the set of admissible past inputs has the form $U = B \ell^p(K_u)$ and that a bound $\|\Gamma_g\|_{\ell^p \rightarrow \ell^\infty} \leq \gamma$ is available as part of the a-priori information, this leads to the following convex relaxation of problem 2.2.1:

**Problem 2.2.2.** Given an unknown plant, the a-priori sets of candidate models, past inputs and noise $(S, U, \mathcal{N})$ and a finite set of samples of the input and output of the plant $(u, y)$ in $[0, N - 1]$, find a model $g \in T(y)$, where:

\[ T(y) \doteq \{ g \in S : y_k - x_k - (T_g u^+)_k \in \mathcal{N} \quad \text{for some } |x|_k \leq \gamma K_u, \]
\[ (k = 0, \ldots, N - 1) \} \quad (2.16) \]

Problem 2.2.2 is solved in the next proposition:
Proposition 2.2.1. Problem 2.2.2 has a solution if the following set of LMIs in $h, x$ is feasible:

\[
M(h) \doteq \begin{bmatrix} KR^{-2} & (T_h^N)^T \\ T_h^N & KR^2 \end{bmatrix} \succeq 0
\]

\[
y - T_u^N P - T_u^N h - x \in N
\]

\[
-\gamma K_u \leq x \leq \gamma K_u
\]

(2.17)

where the last two inequalities should be interpreted in a component-wise sense.

Proof: Straightforward application of Lemma 2.2.2 to the setup given in Fig.2.2 yields the first and the second equations. The third inequality comes from the assumption made during convex relaxation. Thus, the proof is complete.

2.3 Macroeconomic Semi-blind Modeling

The historical time series data containing 1960.1:1996.2 quarter values were manipulated following the Rudebusch and Svensson (RS) model\textsuperscript{12}.

Since measuring the real inflation is not an easy task, we assumed that the measured inflation ($y_k$) is given by

\[
y_k = \pi_k + \eta_k
\]

(2.18)

where $\pi_k$ is the actual inflation and $\eta_k$ is additive noise. From a robust control perspective, the inflation rate will be considered as the output of an LTI system whose input is the FED funds rate. Also, the past input set ($U$), discussed in Section 2.2.1, should be characterized in order to use the semi-blind identification framework. The simplest possible characterization suggests that this set should be represented as

\[
0 \leq i^- \leq i_{max}
\]

(2.19)

where $i_{max}$ is the bound for the historical FED funds rate. Since the actual historical FED funds rate varied in large amounts due to regular and crises modes of the economy, Eqn.2.19 is probably too coarse, possibly yielding very conservative results. However, the quarterly change in the FED funds rate over the period of

\textsuperscript{12}RS model annual percentage inflation and output gap simulations are shown on Fig.2.1. However, the output gap is ignored since this research focuses on the inflation forecasting.
interest is much more uniform and small in terms of actual amounts. Therefore, a much better characterization of the past input set can be obtained when the change in the FED funds rate \( u_k \)

\[ u_k = i_k - i_{k-1} \tag{2.20} \]

is used to identify an operator \( S \) to map the quarterly change of the FED funds rate to inflation \( \pi_{k+1} \). This modification immediately suggests that the operator \( S \) will include an integrator in its parametric portion. Being the last \( a\text{-priori} \) parameter needed before the identification, Fourier analysis of the input-output data for the identification range yielded \( \rho \leq 1.02 \) as the upper bound. With these assumption, the \( a\text{-priori} \) information for the candidate model set,

\[ S = \left\{ H(z) = p_1 \frac{z}{z-1} + G_{np}(z) \mid G_{np}(z) \in \mathcal{BH}_{\infty,\rho}(K) \right\} \tag{2.21} \]

for the noise set,

\[ \mathcal{N} = \{ \eta \in \ell^\infty : |\eta_k| \leq \epsilon \}, \quad \epsilon = 0.2. \tag{2.22} \]

for the past input bound,

\[ \mathcal{U} = \{ u \in \ell^\infty : |u_k| \leq u_{\text{max}} \}, \quad u_{\text{max}} = 0.5 \tag{2.23} \]

and for the unknown initial input(s) conditions set,

\[ x_k = p \cdot i_k + x_{np}^k, \quad \text{where} \quad |x_{np}^k| \leq K \frac{\rho^{-n}}{\rho - 1} 0.5 \tag{2.24} \]

The unknown initial input(s) conditions set given in Eqn.2.24 can be found by recognizing

\[ x_k = p \cdot \sum_{j=-\infty}^{k} (i_j - i_{j-1}) + \underbrace{\Gamma_{np} \cdot u^-}_{\text{where}} \]

\[ \text{where} \quad |(\Gamma_{np} \cdot u^-)_k| = |\sum_{j=-\infty}^{-1} g_{np}^{k-j} u_j| \leq |u_{\text{max}}| \sum_{k+1}^{\infty} |g_j^{np}| \tag{2.25} \]
Since
\[ G^{np} \in B\mathcal{H}_{\infty, \rho}(K) \quad \Rightarrow \quad |g^{np}_n| \leq K\rho^{-n} \]  
(2.26)
holds, it can be substituted for \( g^{np}_j \) term in Eqn.2.25 yielding the Eqn.2.24.

Running the identification algorithm outlined in the Section 2.2.1, on page 18 for \( N = 21 \) historical values of inflation and FED funds rates (from 1961.1 to 1966.1), using \textit{past} inputs as variables, eliminating unobservable/uncontrollable states yields to the following 4\(^{th}\) order discrete-time transfer function
\[
G(z) = \frac{0.522z^4 + 0.208z^3 + 0.184z^2 + 0.276z - 0.024}{z^4 - 0.526z^3 - 0.074z^2 + 0.13z - 0.506}
\]  
(2.27)

Using the time delay property of the z-transform \( \mathcal{Z}^{-1}(z^kG(z)) = g(n + k) \) yields,
\[
\pi_{k+4} = 0.5264\pi_{k+3} + 0.07408\pi_{k+2} - 0.1298\pi_{k+1} + 0.5057\pi_k + 0.5221u_{k+4} + 0.2085u_{k+3} + 0.1842u_{k+2} + 0.2765u_{k+1} - 0.02406u_k
\]  
(2.28)

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure2.4.png}
\caption{The Semi-blind \( \ell_1 \)-Identification Simulation.}
\end{figure}

The forecasting power of this model is illustrated in Figure 2.4, where it was used to estimate the inflation for the entire 1961.1-2003.2 quarters, using past historical data for inflation and changes in FED funds rates as its inputs. As seen,
the model, given in Eqn.2.28, yields a worst–case error of 3.37. For comparison, the RS model has a worst case error of 3.75, (roughly 11% higher) even though it was obtained considering 142 data points. The RS and $\ell_1$ identification inflation forecasting errors(historical-simulation) shown in Fig.2.5.

![Figure 2.5. The RS model and the $\ell_1$–Identification forecasting errors.](image)

It should be noted that the identification performance closely depends on many different parameters chosen. A number of different data length for identification, reduced order systems, different cut-off points for the Fourier transform analysis were used to obtain satisfactory results.

### 2.4 Worst-case Prediction Error Bounds

One of the advantages of the approach outlined in Section 2.2.1 is its ability to provide worst–case bounds on the prediction error. Since the algorithm is interpolatory, it produces a model inside the consistency set $T(y)$. Therefore, since the true system must also belong to the consistency set\textsuperscript{13}, it follows that, given the first $N$ measurements $y_i$, ($i = 0, \ldots, N - 1$) a bound on the worst case prediction

\textsuperscript{13}As long as the a-priori information is correct.
error at \( t = N \) is given by:

\[
|e_N| \leq \sup_{g_1,g_2 \in \mathcal{T}(y)} \left| \left( T_{g_1} - T_{g_2} \right) u^+ + \Gamma_{g_1} (u^-)_1 - \Gamma_{g_2} (u^-)_2 \right|_N
\]

\[
= d[\mathcal{T}(y)]
\]

\[
\leq \sup_y d[\mathcal{T}(y)] = \mathcal{D}(I)
\]

(2.29)

where \( d(.) \) and \( \mathcal{D}(I) \) denote the diameter of the set \( \mathcal{T}(y) \), in the sup–metric, and the diameter of information, respectively. Moreover, since the \( a\text{-priori} \) sets \( (\mathcal{S}, \mathcal{N}) \) are convex and symmetric, with points of symmetry \( g_s = 0 \) and \( \eta_s = 0 \), respectively, by invoking the Lemma B.1.2 on page 84 yields that

\[
\mathcal{D}(I) \leq 2 \sup_{g \in \mathcal{T}(0)} \left| p \cdot i_N + \sum_{j=0}^{N} h_{N-j} \cdot u_j + K \cdot K_u \frac{\rho^{-N+1}}{\rho - 1} \right|
\]

(2.30)

where \( \mathcal{T}(0) \) indicate the set of operators compatible with the zero outcome: \( y_k = 0, k = (0, 1, \ldots, N - 1) \). This leads to the following convex optimization problem:

\[
\max \left| p \cdot i_N + \sum_{j=0}^{N} h_{N-j} \cdot u_j + K \cdot K_u \frac{\rho^{-N+1}}{\rho - 1} \right|
\]

subject to

\[
M(h) = \begin{bmatrix} KR^{-2} & (T_N^h)^T \\ T_N^h & KR^2 \end{bmatrix} \geq 0
\]

\[
\| T_u p + T_u h + x \|_{\infty} \leq \epsilon
\]

\[
|x_j| \leq K \cdot K_u \frac{\rho^{-j}}{\rho - 1}
\]

\[
|u(N)| \leq u_{\max}
\]

\[
|i_N| \leq i_{\max}
\]

\[
|h_N| \leq K \rho^{-N}
\]

The use of this worst case prediction error bound is illustrated by obtaining an estimate of the prediction error at \( N = 22 \), the first historical data point not used in the identification, Eqn.2.31 with \( u_{\max} = 0.6 \) and \( i_{\max} = 6 \), yields \( |e_{22}| \leq 2.37 \).
(actual error is 1.05). The bounds corresponding to $N = 25$ and $N = 30$ are 1.99 and 1.95, showing that, as expected, the error gets smaller as more points are used in the identification.

2.5 Conclusions and Future Research

Semi-blind $\ell_1$-identification is proposed and shown to be capable of explaining time-series data, in the presence of initial conditions.

The proposed identification algorithm can be extended to multi-input, multi-output systems to better represent macroeconomic dynamics whose use for policy analysis is very essential.
Chapter 3

Semi-blind Model (In)validation

The control oriented semi-blind model (in)validation and its application to a general macroeconomic system are explained.

3.1 Model (In)validation

The worst-case, deterministic model (in)validation process is a recently developed engineering practice[26] and it verifies the validity of an already identified model by testing whether it can regenerate the experimental data. Model (in)validation with initial states results in a Bilinear Matrix Inequality which is an NP-hard problem. To circumvent this difficulty, this chapter introduces an LMI based, convex reformulation of the problem. Since it is always possible that all systems can eventually be invalidated for some sequences, model invalidation is emphasized rather than model validation. For an identified model, there is no easy way to show its forecasting power for the future periods. However, model validation presents an opportunity to quantify the forecasting power of an identified model. Model invalidation concept is also consistent with the robust system synthesis framework so that it can be used during optimization. Finally, it easily handles Lucas Critique. In other words, model invalidation test gives a very clear indication as to when the identified model is no longer compatible with the newly measured data and the use of identified model for forecasting and optimization. When implemented,
model invalidation actually computes the minimum norm bound\(^1\) on uncertainty and on disturbance to validate the actual data.

\[
G = (1 + \Delta)G_o
\]  

(3.1)

In general, the robust LTI model for a system with multiplicative uncertainty can be given in Eqn.3.1. Since the set of models consisting of a nominal model \(G_o\) and the bound \(\delta = ||\Delta||_\infty\) on the identification error should belong to the consistency set \(T(y)\), \(G\) is the general representation of the consistency set for \(\Delta \in BH_\infty(\delta)\), i.e., a ball, i.e., a set of models, of \(\delta\) centered at \(G_o\). Note that \(\Delta\) is a stable transfer function that represents the percentual difference among the models in \(T(y)\).

\[\text{Figure 3.1. The model invalidation setup.}\]

Since the system in question may already be in progress as in macroeconomic inflation dynamics, the effect of initial states may be large so that the model validation test may fail. To address this case, following the framework in Fig.3.1, the question of model (in)validation can be presented as

**Problem 3.1.1 (Model Invalidation Problem).** \(\text{Given new data } \{\hat{y}_k, u^+_k\}, \text{ consisting of } N \text{ measurements of the input and output, the nominal model } G_o \text{ and set descriptions } N, BH_\infty(\delta), \text{ and } X_o \text{ of admissible noise, uncertainty and initial conditions, determine if there exists at least one triple } (\eta, \Delta, x_o) \in N \times BH_\infty(\delta) \times X_o \text{ that can reproduce the new data:} \)

\[
\hat{y} = (I + \Delta) \left( T_{G_o} u^+ + T_{G_o}^{nc} x_o + \eta \right)
\]

(3.2)

\(^1\)This bound will be tighter than the one found by calculating an upper bound on the diameter of information.[26]
where $T_{G_0}$, given in Eqn.3.4, and $T_{G_0}^{ic}$ denote the operators that map the current input and system initial conditions of system $G_0$ to its output, respectively.

Affirmative answer to Problem 3.1.1 establishes consistency of the model and gives a powerful tool to forecast future inflation. However, Eqn.3.2 in Problem 3.1.1 contains an $x_o$ term due to initial conditions and when the setup in Figure 3.2 combined with Lemma 2.2.1 is used, this term can not be handled in a convex optimization setup since $T_{G_0}^{ic}$ operator is not known, yielding a bilinear matrix inequality.

### 3.2 Semi-blind Model (In)validation

Since the identified systems are linear time-invariant, there is a linear relationship between the system past input(s) and the system initial input(s) at the system output such that optimization results should be similar whether the system initial conditions or the system past inputs are used, provided that they are mapped to the system output properly. Therefore, addressing the nonconvex problem 3.1.1, the same concept of past input mapping, used during the semi-blind identification, can be utilized, i.e., the contribution of the $T_{G_0}^{ic}x_o$ term can be represented by some unknown input $u^- \in U$ acting in $[-T, -1]$ for some $T > 0$, which is depending on the model order and by using the associated Hankel operator.

Let $g[n]$ be the impulse response of the discrete-time nominal model $G_o(z)$, $u[n]$ be the arbitrary input sequence, $y[n]$ be the corresponding outputs. Following the convolution analysis for linear, time-invariant discrete-time systems,

$$ y[n] = \sum_{k=0}^{N} g[k].u[n - k] \quad (3.3) $$

can be written. Notice that the convolution sum given in Eqn.3.3 can not handle the system initial conditions. However, the simulation can still start from $k = 0$ and the past inputs($u_k^-$) from $(k = -1, \ldots, -T)$ mapping through a Hankel operator for the corresponding initial conditions of the system can be included. The system output for this approach can be written in Toeplitz matrix form for
clarity as below,

\[
\begin{bmatrix}
  y[0] \\
  y[1] \\
  \vdots \\
  y[N-1]
\end{bmatrix}
= \begin{bmatrix}
  g[0] & 0 & \cdots & 0 \\
  g[1] & g[0] & \cdots & 0 \\
  \vdots & \vdots & \ddots & \vdots \\
  g[N-1] & g[N-2] & \cdots & g[0]
\end{bmatrix}
\begin{bmatrix}
  u[0] \\
  u[1] \\
  \vdots \\
  u[N-1]
\end{bmatrix}
\]

\[
+ \begin{bmatrix}
  g[1] & g[2] & \cdots & g[T] \\
  g[2] & g[3] & \cdots & g[T+1] \\
  \vdots & \vdots & \ddots & \vdots \\
  g[N+1] & g[N+2] & \cdots & g[T+N]
\end{bmatrix}
\begin{bmatrix}
  u[-1] \\
  \vdots \\
  u[-T]
\end{bmatrix}
\]

\[(3.4)\]

It should be noted that for a \(T^{th}\)-order system, it needs \(T\) past input conditions and for any controllable system with constraints on the control input may take arbitrarily high number of steps to reach the final state, depending on the type of the constraint. With the new approach, the modified validation problem can be given in Fig.3.2 and formulated as

\[
\hat{y} = (I + \Delta) \left( T_{G_o} u^+ + H_{G_o} u^- \right) + \eta \quad (3.5)
\]

**Figure 3.2.** The semi-blind model invalidation setup.

**Problem 3.2.1 (Semi-blind Model Validation Problem).** Given new data \(\{\hat{y}_k, u^+_k\}\), consisting of \(N\) measurements of the quarterly inflation and the change in FED short term rate, the nominal model \(G_o\) and set descriptions \(\mathcal{N}, \mathcal{X}_o\) and \(\mathcal{BH}_\infty(\delta)\) of admissible noise, uncertainty and initial conditions, determine if there exists at least one triple \((\eta, \Delta, u^-) \in \mathcal{N} \times \mathcal{BH}_\infty(\delta) \times \mathcal{U}\) that can reproduce the new data:
where \( T_{G_o} \) and \( H_{G_o} \) denote the operators that map the current and past inputs of system \( G_o \) to its output, given in Eqn.3.4.

By using the results of model (in)validation[49] and the notation of \( T^N_\xi \) for the corresponding \( (N \times N) \) Toeplitz matrix for \( \xi \) variable and \( T^N_\omega \) for other variables accordingly, the existence of an uncertainty block \( \Delta \) in Fig.3.2 such that \( ||\Delta||_\infty \leq \delta \) is equivalent to,

\[
(T^N_\xi)^T (T^N_\xi) \leq \delta^2 (T^N_\omega)^T (T^N_\omega) \tag{3.6}
\]

where \( N \) is the length of the data points. By using the Fig.3.2, the nominal model output in terms of the inputs

\[
T^N_\omega \doteq T^N_{G_o} T^N_u + H^N_{G_o} T^N_u \tag{3.7}
\]

and the uncertainty block output in terms of the other variables

\[
T^N_\xi = T^N_y - T^N_\omega - T^N_\eta \tag{3.8}
\]

can be written. Now, replacing the \( T^N_\xi \) variable in Eqn.3.8 with the one in Eqn.3.6 and rearranging,

\[
\delta^2 (T^N_\omega)^T (T^N_\omega) - \left( (T^N_y)^T - (T^N_\omega)^T - (T^N_\eta)^T \right) \left( T^N_y - T^N_\omega - T^N_\eta \right) \geq 0 \tag{3.9}
\]

can be obtained. Expanding Eqn.3.9 yields:

\[
\delta^2 (T^N_\omega)^T (T^N_\omega) - \left( (T^N_y)^T - (T^N_\omega)^T - (T^N_\eta)^T \right) (T^N_y - T^N_\omega - T^N_\eta) \geq 0 \tag{3.10}
\]

and

\[
\delta^2 (T^N_\omega)^T (T^N_\omega) - (T^N_y)^T (T^N_y - T^N_\omega - T^N_\eta) + (T^N_\omega)^T (T^N_y - T^N_\omega - T^N_\eta) + (T^N_\eta)^T (T^N_y - T^N_\omega - T^N_\eta) \geq 0 \tag{3.11}
\]

can be written. Now, multiplying both sides by \(-1\) and grouping the common
terms,

\[(1 - \delta^2) (T_N^\omega)^T (T_N^\omega) + (T_N^y)^T (T_N^y) - (T_N^\eta)^T (T_N^\eta)\]
\[- (T_N^y)^T (T_N^\eta) - (T_N^\omega)^T (T_N^\eta) + (T_N^\omega)^T (T_N^\omega)\]
\[- (T_N^\eta)^T (T_N^y) + (T_N^\eta)^T (T_N^y) + (T_N^\omega)^T (T_N^\omega) \leq 0\]  \hspace{1cm} (3.12)

can be obtained. However, Eqn.3.12 has several terms that lead to a bilinear matrix inequality (BMI), not joint convexity in all variables. For example, when \((T_N^\eta)^T (T_N^\omega)\) term is taken in the last line of Eqn.3.12 and is substituted by \(T_N^\omega = T_N^{G_0} (u^+ + H_N^{G_0} u^-)\), cross-terms² like \((T_N^\eta)^T H_N^{G_0} T_N^{G_0}\) are obtained, which make the inequality a bilinear matrix inequality (BMI).

To overcome this jointly nonconvex problem, we will use a deterministic convex relaxation similar to that proposed in [50] where the noise input is moved ahead of the uncertainty block as shown in Fig.3.3.

### 3.3 A Deterministic Convex Relaxation

![Figure 3.3. The semi-blind model invalidation convex relaxation setup.](image)

The semi-blind model invalidation convex relaxation setup is given in Fig.3.3 whose only difference with the original setup in Fig.3.2 is the point where noise enters the system, the noise is assumed to be affected by the uncertainty. For any \((u^-, \tilde{\eta}, \delta)\) triple satisfying

\[y = (I + \Delta) (T_{G_0} u^+ + H_{G_0} u^- + \tilde{\eta})\]  \hspace{1cm} (3.13)

²Note that noise and the past inputs are variables to perform model invalidation through uncertainty minimization.
with \( \|\tilde{\eta}\|_2 \leq \bar{\epsilon} \) and \( \|\Delta\|_{\infty} \leq \delta \), then the \((u^-, \eta, \delta)\) triple with \( \eta = (1 + \Delta)\tilde{\eta} \) satisfies the original setup in Fig.3.2. Therefore, the model invalidation process can be done by searching for a solution of the problem in Fig.3.3, with the noise bound \( \bar{\epsilon} = \frac{\epsilon_1}{1 + \|\Delta\|_{\infty}} \). Although the noise bound is scaled by \( \frac{1}{1 + \|\Delta\|_{\infty}} \) in the new setup, since the uncertainty bound is expected to be small, then the new setup is not too conservative. Thus, the semi-blind model invalidation with convex relaxation result can be stated as

**Proposition 3.3.1.** There exists a feasible triple \((\eta, \Delta, u^-) \in \mathcal{N} \times \mathcal{B}H_{\infty}(\delta) \times \mathcal{U}\) that satisfies Eqn.3.13 if and only if there exists at least one pair of finite sequences \( u^- = \{u_{-1}^-, u_{-2}^-, \ldots, u_{-(T-1)}^-\} \in \mathcal{U} \) and \( \eta \in \mathcal{N} \) such that the following LMI holds:

\[
A(\eta, \delta, u^-) = \begin{bmatrix}
X(u^-, \eta) & Y^T(u^-, \eta) \\
Y(u^-, \eta) & (\delta^2 - 1)^{-1}I
\end{bmatrix} \leq 0 \quad (3.14)
\]

where

\[
Y(u^-, \eta) = T_N^N T_N u^+ + H_N^N T_N u^- + T_N^\eta \\
X(u^-, \eta) = (T_N^y y) - T_N^y - (T_N^y)^T \cdot Y(u^-, \eta) - Y^T(u^-, \eta) \cdot T_N^y
\]

**Proof:** Using the setup in Fig.3.3 and the model invalidation results[49],

\[
(T_\xi^N)^T (T_\xi^N) \leq \delta^2 (T_\omega^N)^T (T_\omega^N) \quad (3.15)
\]

is written. Substituting \( T_\xi^N = T_N^y - T_N^\omega \) in Eqn.3.15,

\[
((T_N^y)^T - (T_N^\omega)^T)((T_N^y)^T - (T_N^\omega)^T) \leq \delta^2 (T_N^\omega)^T (T_N^\omega) \quad (3.16)
\]

Rearranging the Eqn.3.16 yields,

\[
(T_N^y)^T (T_N^y) - (T_N^\omega)^T (T_N^\omega) - (T_N^\omega)^T (T_N^\omega) \leq 0 \quad (3.17)
\]
Applying the Schur’s complement property\(^3\) to Eqn.3.17

\[
\begin{bmatrix}
(T_y^N)^T (T_y^N) - (T_y^N)^T (T_\omega^N) - (T_\omega^N)^T (T_y^N) \\
(T_\omega^N) - (\delta^2 - 1)^{-1} I
\end{bmatrix} \leq 0
\]

(3.18)

can be written. Recognizing \(T_z^N = T_{Go}^N T_{u^+}^N + H_{Go}^N T_{u^-}^N\) from the setup in Fig.3.3 and substituting it into \(T_\omega^N = T_z^N + T_\eta^N\) yields

\[
T_\omega^N = \underbrace{T_{Go}^N T_{u^+}^N + H_{Go}^N T_{u^-}^N + T_\eta^N}_{T_z^N}
\]

(3.19)

which yields after substituting in Eqn.3.18,

\[
\begin{bmatrix}
(T_y^N)^T (T_y^N) - (T_y^N)^T (T_{Go}^N T_{u^+}^N + H_{Go}^N T_{u^-}^N + T_\eta^N) \\
-(T_{Go}^N T_{u^+}^N + H_{Go}^N T_{u^-}^N + T_\eta^N)^T (T_y^N) \\
(T_{Go}^N T_{u^+}^N + H_{Go}^N T_{u^-}^N + T_\eta^N) - (\delta^2 - 1)^{-1} I
\end{bmatrix} \leq 0
\]

(3.20)

By comparing each term in Proposition.3.3.1 and in Eqn.3.20, the proof is completed.

\(^3\)Schur’s complement property simply says that for given \(X \in H^{n \times n}, Y \in H^{m \times m}\), where \(H^{n \times n} := \{M \in C^{n \times n}; M^* = M\}\) is the set of complex Hermitian \((n \times n)\) matrix, \(Z \in C^{n \times m}\), where \(C\) is the complex number set, the matrix inequalities below are equivalent;

\[
\begin{bmatrix}
X \\
Z^*
\end{bmatrix} > 0 \iff Y > 0 \text{ and } X - ZY^{-1}Z^* > 0.
\]
To show the LMI approach, rearranging the variables in Eqn.3.20 yields

\[
\begin{bmatrix}
(T_N^y)^T (T_N^y) - (T_N^y)^T (T_N^G)(T_N^u^+) & (T_N^u^+)^T (T_N^G)^T + (T_N^u^-)^T (H_N^G)^T \\
-(T_N^y)^T (H_N^G)(T_N^u^-) - (T_N^y)^T (T_N^H) & + (T_N^H)^T \\
-(T_N^u^+)^T (T_N^G)^T (T_N^y) - (T_N^u^-)^T (H_N^G)^T (T_N^y) & - (T_N^H)^T (T_N^H)
\end{bmatrix} \leq 0
\]

\[(3.21)\]

Letting

\[
A = (T_N^y)^T (T_N^y) - (T_N^y)^T (T_N^G)(T_N^u^+) - (T_N^u^+)^T (T_N^G)^T (T_N^y)
\]

\[
B = (T_N^y)^T (H_N^G)
\]

\[
X_1 = T_N^H
\]

\[
X_2 = (\delta^2 - 1)^{-1} I
\]

\[
X_3 = T_N^u^-
\]

and noticing that \(A, B\) are a-priori known while \(X_1, X_2\) and \(X_3\) are variables,

\[
\begin{bmatrix}
A - B \cdot X_3 - (T_N^y)^T \cdot X_1 & (T_N^u^+)^T \cdot (T_N^G)^T + X_3 \cdot (H_N^G)^T + X_1^T \\
-X_3 \cdot H_N^G \cdot T_N^y & -X_1^T \cdot T_N^y
\cdot T_N^G \cdot T_N^u^+ + H_N^G \cdot X_3 + X_1 & X_2
\end{bmatrix} \leq 0
\]

\[(3.22)\]

can be obtained. Since all the variables enter the matrix expressions linearly, the matrix inequality in Eqn.3.22 is a linear matrix inequality (LMI) and can be solved by using Matlab LMI toolbox\(^4\).

### 3.4 Macroeconomic Semi-blind Model Invalidation

After a macroeconomic model is identified, it should be validated with new data so that its practical use will yield plausible policy choices. Since model invalidation

\(^4\)The Matlab LMI files will be given in the appendix A on page 81 in the final copy.
approach determines clearly that when a model is no longer valid or compatible for some new data points, this approach can be used during policy optimization or forecasting frameworks immune to Lucas Critique.

The macroeconomic system semi-blind model invalidation setup can be depicted as shown in Fig.3.2, where $u^-$ and $u^+$ denote the past and current FED funds rates, respectively, $y$ denotes the annual inflation, the nominal model $G_o$, which was found during the mixed parametric/nonparametric $\ell_1$ identification algorithm and given in Eqn.2.27, the $\eta$ noise bound is taken to be $|\eta|=0.2$ to represent inflation estimation irregularities\(^5\).

Using the LMI variables related to the past input variable($X_3$), the noise variable($X_1$) and the uncertainty bound($X_2$) in the Eqn.3.22, the semi-blind model validation with convex relaxation approach calculations were performed by using the Matlab LMI toolbox. For comparison purposes, the new model invalidation approach was applied to both the $\ell_1$-identification and RS models for the whole data range. The invalidation results are given in Table.3.1 and Table.3.2, which clearly shows that the $\ell_1$-identification model performs better than the RS model such that the uncertainty bound is up to 30% to 50% better than what RS model generated.

\begin{table}[h]
\centering
\caption{Model Invalidation Results for the same prediction error bound}
\begin{tabular}{|c|c|c|c|}
\hline
Data range & Prediction error & $\|\Delta\|_{\infty}$ & $\|\Delta_{RS}\|_{\infty}$ \\
\hline
5-170 & 1.7 & 0.5 & 0.63 \\
5-170 & 2 & 0.3 & 0.4575 \\
5-170 & 3.372 & 0.000884 & 0.037 \\
\hline
\end{tabular}
\end{table}

\subsection{3.5 Conclusions and Future Research}

In this chapter we have shown that policy optimization coupled with semi-blind model invalidation approach can determine when a system is no longer valid for

\(^5\)Since the historical US inflation is less than 20\%, this noise bound corresponds to 1\% inflation estimation error. However, when the inflation is low, i.e., around 2\%, this noise bound now corresponds to 10\% inflation estimation error.
Table 3.2. Model Invalidation Results for the same uncertainty bound

| Data range | Uncertainty bound | $|e|$ | $|e_{RS}|$ |
|------------|-------------------|------|-----------|
| 5-170      | 0.05              | 3.0069 | 3.3066 |
| 5-170      | 0.2               | 2.2285 | 2.7085 |
| 5-170      | 0.5               | 1.7227 | 1.9123 |
| 5-170      | 0.9999            | 1.2853 | 1.3483 |

the underlying dynamics during the optimization and thus is immune to Lucas critique.

Semi-blind model invalidation concept can be applied to $\ell_\infty$-to-$\ell_\infty$ uncertainty or other uncertainty cases to reflect the real-life situations.
Chapter 4

Linear Fractional Transformations (LFT)

In this chapter, the linear fractional transformations (LFT), robust controller synthesis and an optimal controller design of a general macroeconomic stabilization framework are explained and possible FED policy preferences are interpreted.

4.1 Linear Fractional Transformations

LFT framework\[54\] is a very powerful approach to represent uncertainty in system modeling and appeals to the increasing number of researchers in the robust control community. The main idea is to group known and unknown quantities separately in a feedback configuration and set bounds for the unknown quantities. Then, algebra and matrix properties are used to simplify the block representation and to carry out optimization steps. The fundamental benefits of using LFTs in robust control design are the facts that LFTs convert uncertainties at the component level to structured uncertainties at the interconnection level and that interconnections of several LFTs are also LFTs.

To introduce the basic LFT framework, consider two matrices\(^1\) \(G\) and \(\Delta\) representing two different systems shown in the Fig.4.1 and assume that \(G\) matrix can be partitioned as \(G = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} \).

\(^1\)\(G\) and \(\Delta\) matrices will include the known quantities and uncertainties, respectively.
Either upper or the lower loop can be closed through $\Delta$ block, yielding a name of upper or lower linear fractional representation. However, with the representation and relationships shown in the Fig.4.1, it is obvious that the upper loop is closed. When the upper loop of $M$ is closed with $\Delta$, the linear fractional transformation of $M$ by $\Delta$ can be depicted in Fig.4.2.

From Fig.4.2, it is possible to find a transfer function between $u_2$ and $y_2$ by eliminating $u_1$ and $y_1$ through $\Delta$ relationship. The $u_1 = \Delta y_1$ can be substituted in

$$y_1 = G_{11} u_1 + G_{12} u_2$$
$$y_2 = G_{21} u_1 + G_{22} u_2$$

yielding

$$y_1 = G_{11} \Delta y_1 + G_{12} u_2$$
$$y_2 = G_{21} \Delta y_1 + G_{22} u_2$$
grouping the same terms

\[(I - G_{11}.\Delta)y_1 = G_{12}.u_2\]
\[\downarrow\]
\[y_1 = (I - G_{11}.\Delta)^{-1}.G_{12}.u_2\]
\[y_2 = G_{21}.\Delta.y_1 + G_{22}.u_2\]

and substituting \(y_1\) expression in the second equality in the Eqn.4.3

\[y_2 = G_{21}.\Delta.(I - G_{11}.\Delta)^{-1}.G_{12}.u_2 + G_{22}.u_2\]
\[\downarrow\]
\[y_2 = \left(G_{21}.\Delta.(I - G_{11}.\Delta)^{-1}.G_{12} + G_{22}\right).u_2\]
\[\downarrow\]
\[y_2 = F_U(M, \Delta).u_2\]

(4.4)

can be obtained. The \(F_U(M, \Delta) = (G_{21}.\Delta.(I - G_{11}.\Delta)^{-1}.G_{12} + G_{22})\) is the upper\(^2\) LFT representation of the system given in Fig.4.2.

To show the representation of uncertainty by LFTs, consider an uncertain parameter \(K\) such that \(1 \leq K \leq 2\). Obviously, \(K = 1.5 + 0.5\delta_k\), where \(\delta_k \in [-1, 1]\) can also be written for the same uncertain parameter. Assuming that the upper LFT representation is used, \(K = 1.5 + 0.5\delta_k\) expression can be compared with \(F_U(M, \delta_k) = (G_{21}.\delta.(I - G_{11}.\delta)^{-1}.G_{12} + G_{22})\) and the appropriate terms can be determined. From this example, \(G_{22} = 1.5\), \(G_{21} = 0.5\), \(G_{11} = 0\) and \(G_{12} = 1\) can be obtained such that

\[K = F_U\left(\begin{bmatrix} 0 & 1 \\ 0.5 & 1.5 \end{bmatrix}, \delta_k\right)\]

(4.5)

and the upper LFT representation is given in Fig.4.3. Finally, the LFT representation in Fig.4.3 can be replaced wherever \(K\) parameter exists in a block diagram.

If \(\frac{1}{K}\) is in the block diagram, LFT approach can still be used due to the fact that the inverse of an LFT is another LFT in terms of \(\delta\). For the previous example,

\(^2\)If the lower loop is closed through appropriate \(\Delta\) block, the lower LFT representation \(F_L(M, \Delta) = (G_{11} + G_{12}.\Delta.(I - G_{22}.\Delta)^{-1}.G_{21})\) such that \(y_1 = F_L(M, \Delta).u_1\) can be shown easily by following the same steps above.
The inverse of $K = 1.5 + 0.5\delta_k$ is

\[
\frac{1}{K} = \frac{1}{1.5 + 0.5\delta_k} = \frac{1}{1.5} + \frac{\delta_k}{1.5}\frac{1}{1.5} = -\frac{\delta_k}{3} \left( 1 - \left(\frac{-1}{3}\right)\delta_k \right)^{-1} \frac{1}{1.5} + \frac{1}{1.5}
\]

\[
\equiv (G_{21}.\Delta.(I - G_{11}.\Delta)^{-1}.G_{12} + G_{22})
\]

\[
= F_U \left( \begin{bmatrix} -\frac{1}{3} & \frac{1}{1.5} \\ -\frac{1}{3} & \frac{1}{1.5} \end{bmatrix}, \delta_k \right)
\]

(4.6)

In general, the inverse of an LFT can be calculated by using the matrix inversion lemma. To show this, for a system $G = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix}$ and the uncertainty block $\Delta$ and a given upper LFT representation $F_U(G, \Delta)$

\[
F_U(G, \Delta)^{-1} = F_U \left( \begin{bmatrix} G_{11} - G_{12}G_{22}^{-1}G_{21} & G_{12}G_{22}^{-1} \\ -G_{22}^{-1}G_{21} & G_{22}^{-1} \end{bmatrix}, \Delta \right)
\]

(4.7)

can be written for the inverse of the LFT.

**Definition 4.1.1 (Well-posedness).** [54] Given a system $G = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix}$ and an uncertainty $\Delta$ such that $B_\Delta \doteq \{ \Delta \in L(L_2) : ||\Delta|| \leq 1 \}$, the LFT system in Figure D.1 is robustly well-posed if

\[
\det(I - G_{11}\Delta) \neq 0
\]

holds. In other words, $I - G_{11}\Delta$ is nonsingular for all $\Delta \in \Delta$. Well-posedness guarantees that all internal vectors are unique for any external vectors.

**Corollary 4.1.1 (Normalizing uncertainty).** [54] If the uncertainty description
in Definition 4.1.1 is given as \( B_\Delta = \{ \Delta \in \mathcal{L}(L_2) : ||\Delta|| \leq \beta \} \), the corresponding well-posedness test would be \( ||G_{11}|| < 1/\beta \).

4.2 Robust Synthesis Problem

Robust synthesis problem involves designing a controller to stabilize an LFT structure and to achieve given performance objectives such as minimizing some transfer function norms, etc. If the performance objective for an LFT diagram is satisfied for the nominal plant, it is said that the nominal performance is obtained. This is done by checking the norm of the performance objective. If the stability of the LFT system is satisfied for all possible plants for a given plant family, it is said that the robust stability is obtained. By using the small-gain theorem, robust stability condition is equal to a norm test of the weighted nominal closed-loop transfer function. If, for all possible plant models for a given plant family, the LFT system is stable and, at the same time, the performance objective is satisfied, it is said that the robust performance is obtained. The robust performance property is determined by using a structured singular value\((\mu)\) test of the nominal closed-loop transfer function.

It is plausible to expect that the performance and stability of an LFT system is closely dependent to the structure of uncertainty. By manipulating the structure of the uncertainty block \( \Delta \), optimization solution can be obtained in less and less conservative fashion;

- If the uncertainty block \( \Delta \) contains all nonzero elements, i.e., there is no structural relationship among different uncertainty components, this problem yields a well-known \( H_\infty \) problem whose solution is also well-known\(^3\). Although the full block uncertainty description is appealing in terms of mathematical tractability, the lack of information about the structure of the uncertainty usually causes too conservative results, which is not desired in many engineering applications.

- If the uncertainty block \( \Delta \) has some structure such as the ones coming from

\(^3\)In \( H_\infty \) framework, the uncertainty and the performance criteria are given in terms of infinity norms and this section uses this criteria. The solution of \( H_\infty \) is given in different forms such as using algebraic Riccati equations\(^59\), using LMI approach\(^60\).
LFT framework, this problem yields a $\mu$-synthesis problem whose solution is more involved at the benefit of achieving better results.

![Figure 4.4. The Robust Control framework.](image)

The LFT system in Fig.4.4 includes $G$ which contains the nominal plant model and performance and uncertainty weighting functions, $\Delta$ which denotes the norm-bounded uncertainty and $K$ which represents the controller. While $z$, $w$, $u$ and $y$ signals are self-explanatory, $d$ and $e$ signals represent the input and output signals, respectively, to measure performance specifications. Since the uncertainty block $\Delta_U$ is assumed to have diagonal structure

$$
\Delta_U = \begin{bmatrix}
\Delta_1 & & \\
& \cdots & \\
& & \Delta_F \\
\delta_1 I_{r_1} & & \\
& \cdots & \\
& & \delta_s I_{r_s}
\end{bmatrix}
$$

(4.8)

this controller design will be of a $\mu$-synthesis problem. This uncertainty structure is also compatible with the macroeconomic system LFT framework studied in Section 4.3. The uncertainty block is modified by adding another fictitious uncertainty block to reflect the performance objectives condition and the augmented
uncertainty structure(Δ) can be obtained and given in Eqn.4.9.

\[ \Delta = \begin{bmatrix} \Delta_U \\ \Delta_P \end{bmatrix} \text{ where } \Delta_U \in \Delta, \Delta_P \in \mathbb{C}^{m_d \times m_e} \] (4.9)

Whenever the uncertainty block is structured, a structured singular value(μ) can be used to obtain robustness properties of an LFT diagram. The structured singular value is defined for a specific structure depending on uncertainty and performance objectives of a problem and for a specific system matrix. The formal definition can be given as follows;

**Definition 4.2.1 (Structured Singular Value).** [54] Given a system G and an uncertainty structure Δ, the structured singular value, \( \mu_\Delta(G) \), is defined as

\[
\mu_\Delta(G) \doteq \frac{1}{\min\{\sigma(\Delta) : \Delta \in \Delta, \det(I - G\Delta) = 0\}}
\] (4.10)

Note that \( \mu_\Delta(G) = 0 \) if no \( \Delta \in \Delta \) satisfies \( \det(I - G\Delta) = 0 \).

By using the special case of the main-loop theorem, \( \mu \)-synthesis can be defined as

**Theorem 4.2.1 (\( \mu \)-Synthesis\(^4\)).** [54] For the system given in Fig.4.4 and the uncertainty \( \Delta_U \in \Delta \), the system is stable and has \( \overline{\sigma}(T_{de}) < 1 \), if and only if \( F_l(G,K) \) is stable and

\[
\max_\omega \mu_\Delta(F_l(G,K)(j\omega)) < 1
\]

and the robust performance problem for one uncertainty block(\( \Delta_U \)) can be tackled as the robust stability problem for two uncertainty blocks, one for the uncertainty and one for the performance objective, by taking the augmented uncertainty structure given in Eqn.4.9.

Although the exact calculation of \( \mu_\Delta(.) \) is not known, the upper and lower bounds have been developed. When the uncertainty block contains all nonzero

---

\(^4\)This theorem is a different version of the Main-loop theorem, and simply says that if a \( \mu \)-test can be used to determine a particular property of a matrix, then this \( \mu \)-test can be generalized to determine whether the property is robust to structured perturbations.
elements, \[
\max_{\Delta \in \Delta} \rho(\Delta G) = \mu_\Delta(G)
\]
where \( \rho(.) = \lambda_{\text{max}}(.) \) is the spectral radius of a matrix, can be written. Although there is no known algorithm to calculate the global maximum, algorithms exist to calculate the local maximums by using the maximum modules theorem\cite{55}. On the other hand, when the uncertainty is a single full block,

\[
\mu_\Delta(G) = \bar{\sigma}(G)
\]

can be written for the upper bound. Therefore, by using the definition of \( \mu \) and the full block case, the range of the structured singular value

\[
\max_{\Delta \in \Delta} \rho(\Delta G) = \mu_\Delta(G) \leq \bar{\sigma}(G)
\]

is obtained. However, the upper bound and actual \( \mu_\Delta(.) \) can be much smaller depending on the uncertainty structure, i.e., the gap between \( \mu_\Delta(.) \) and \( \bar{\sigma}(.) \) could be arbitrarily large. To minimize this gap, the upper bound is minimized by using transformations that affect \( \bar{\sigma}(.) \) but don’t affect \( \mu_\Delta(.) \). This transformation is done by picking a hermitian positive definite \( D \in \mathbb{D} = \mathbb{C}^{n\times n} \) matrix such that \( D.\Delta = \Delta.D \) holds\footnote{This commutative property is satisfied only when the structure of \( D \) is identical with the structure of \( \Delta \).\cite{58}}. This yields to \( \mu_\Delta(G) = \mu_\Delta(DGD^{-1}) \). Then, the new bounds can be written as

\[
\max_{\Delta \in \Delta} \rho(\Delta G) = \mu_\Delta(G) \leq \inf_{D \in \mathbb{D}} \sigma(DGD^{-1})
\]

The good convexity properties\footnote{The \( \bar{\sigma}(D^{1/2}GD^{-1/2}) \) function possesses a convexity property such that, for compatible \( G \) and \( D \) matrices and \( \beta > 0 \), \( \bar{\sigma}(D^{1/2}GD^{-1/2}) < \beta \Leftrightarrow \lambda_{\text{max}}(G^*DG - \beta^2D) < 0 \) holds. By using linear algebra, it can be shown that this is also valid for when \( D^{1/2} \) is added without changing the upper bound.} enable us to do the minimization by using Linear Matrix Inequalities. Moreover, the upper bound is exactly equal to \( \mu_\Delta(.) \) when \( 2S + F \leq 3 \), where \( S \) is the number of scalar blocks and the \( F \) is the number of full blocks in the uncertainty set.

Now that all the necessary tools are explained, the \( \mu \)-synthesis problem for the
The LFT diagram given in Fig.4.4 can be stated as follows;

**Problem 4.2.1 (µ-Synthesis Problem).** \[54\] Minimize the peak value of \( \mu_\Delta(\cdot) \) of the \( F_i(G, K) \) transfer function of the closed loop system, over all stabilizing controllers \( K \). Stated differently,

\[
\min_{k_{\text{stabilizing}}} \max_\omega \mu_\Delta(F_i(G, K)(j\omega)) \tag{4.13}
\]

The method to calculate the structural singular value can be used in Eqn.4.13 to obtain

\[
\min_{k_{\text{stabilizing}}} \max_\omega \mu_\Delta(F_i(G, K)(j\omega)) \approx \max_\omega \sigma(D_\omega F_i(G, K)(j\omega)D^{-1}_\omega) \tag{4.14}
\]

Since \( D \)-scaling is chosen independently for every frequency \( \omega \), and \( \max_\omega \sigma(.) = ||.||_\infty \), yields

\[
\min_{k_{\text{stabilizing}}} \min_{D_\omega \in D} \max_\omega \sigma(D_\omega F_i(G, K)(j\omega)D^{-1}_\omega) \tag{4.15}
\]

Since the elements of \( D \)-scaling matrix can take nonzero complex values without changing the upper bound, \( D \) matrix can be any real, rational, stable and minimum-phase transfer function. Thus, modifying the Eqn.4.15 for the upper bound

\[
\min_{k_{\text{stabilizing}}} \min_{\tilde{D}(\cdot) \in D} ||\tilde{D}F_i(G, K)\tilde{D}^{-1}||_\infty \tag{4.16}
\]

can be written. This upper bound is solved by 'D-K Iteration' approach\(^7\);

---

\(^7\)The drawbacks of D-K iteration are that the \( \mu_\Delta(.) \) is approximated by its upper bound and that the iteration may get stuck at a local minimum.
• \( \tilde{D}(s) \) fixed: For a given (fixed) real, rational, stable, minimum-phase \( \tilde{D}(s) \), this is absorbed in the nominal system matrix \( G \), as shown in Fig.4.5. By using the overall system \( G_D \), it is obvious that \( K \) stabilizes \( G_D \) if and only if it stabilizes \( G \). Then the Eqn.4.16 becomes

\[
\begin{align*}
\min_{K \text{ stabilizing}} \| \tilde{D} F_l(G, K) \tilde{D}^{-1} \|_\infty = \min_{K \text{ stabilizing}} \| F_l(G_D, K) \|_\infty
\end{align*}
\]

which is an \( H_\infty \) problem and can be solved easily in different forms[59, 60].

• \( K \) fixed: After a controller \( K \) is designed, this controller is held fixed and

\[
\min_{D(s) \in D} \sigma(D(w) F_l(G, K)(j\omega) D^{-1}(w))
\]

minimization problem is solved. This optimization fits all frequency responses of the elements of the transfer function matrix \( D(s) \) by proper, stable, minimum-phase transfer functions.

• Iterate: Once the new \( D(s) \) is found, it is absorbed into the generalized plant, as shown in Fig.4.5 and new iteration starts. This process is done until the \( \mu \) value falls below 1, which implies robust performance of the LFT system.

So far, the stability and performance measures are given in the same norms. If different norms are used, the problem turns into a multi-objective problem, which is still a developing research area. However, there are some results available for different setups[56, 57].

### 4.3 Macroeconomic Stabilization Problems

Assuming a dynamic relationship between the Federal Reserve Bank(FED) short-term interest rates and the annual inflation, the model uncertainty effect on the inflation stabilization can be overcome by using robust control optimization framework, explained in Section 4.2. To improve the overall performance of the economic system, different uncertainty properties, represented by different frequency characteristics and relative weight functions with respect to other performance objective
Macroeconomic stabilization problems can be initiated by establishing a model between input and output variables. The $G_N[z]$ transfer function model that was found in the model identification and validation chapters 2 and 3;

\[
G_N[z] = \frac{Y[z]}{U[z]} = \frac{0.522z^4 + 0.208z^3 + 0.184z^2 + 0.276z - 0.024}{z^4 - 0.526z^3 - 0.074z^2 + 0.13z - 0.506} \tag{4.19}
\]

where $Y[z]$ is the annual inflation output variable and $U[z]$ is the FED short-term interest rate change, found in Section 2.3, is used. Since the $\infty$-norm is transferable in continuous and discrete-times and this conversion is a common practice, the 'bilinear' transformation\(^8\) is used to obtain $G_N(s)$, the continuous-time equivalent of $G[z]$,

\[
G_N(s) = \frac{0.2426s^4 + 5.686s^3 + 12.84s^2 + 20.08s + 22.86}{s^4 + 17.98s^3 + 15.25s^2 + 46.2s + 0.4706} \tag{4.20}
\]

and it can be expanded to $G(s) = G_N(s)(1 + \Delta(s))$, where $||\Delta||_\infty \leq 1$ to denote the multiplicative uncertainty.

As explained before, robust performance implies that the system is stable for given uncertainty set and, at the same time, the performance objectives are satisfied for all possible uncertainty levels in a pre-defined uncertainty set. The general loop block diagram including the nominal plant, a controller, a multiplicative uncertainty, weight transfer functions, etc. can be given as in Fig.4.6.

The robust performance objectives should be determined first to assure placing the correct weighting functions at the correct locations. In general, performance objectives are weighted transfer functions which are made small by designing a controller. Weighting functions can be seen as frequency-dependent scaling of the corresponding transfer functions or signals. The inflation-targeting problems try to minimize the inflation rates with the optimum monetary policy while maximizing the GDP output. Since the model contains only the inflation dynamic, minimum tracking error due to disturbance input or exogenous shocks at the inflation output.

\(^8\)The bilinear transformation is done by substituting $z = e^{sT} \approx \frac{1 + sT}{1 - sT}$ or $s \approx \frac{2(z-1)}{T(z+1)}$. 
is taken one of the performance objective and

\[ W_e(s) = 150 \frac{s + 1}{0.001 + 1} \]  

(weighting function, a low-pass function, is chosen to emphasize on the steady-state error which is assumed to be the main concern and is placed on the path of the weighted error output, \( e(s) \)). However, there is no exact relationship between the minimum steady-state error and the weighting function values and the exact weight function selections for different performance objectives is still an open research area. Therefore, weight selection process should be iterated to reflect the real-life situations.

\( W_d(s) \) weight transfer function can be placed on the disturbance input path so that the disturbance characteristics can be defined by properly adjusting the \( W_d(s) \) function. For example, if the noise is the main concern for disturbance, a high-pass filter should be used. On the other hand, if sustained bounded shocks are of the main concern, the \( W_d(s) \) should be adjusted properly to reflect the disturbance spectrum. For simplicity and brevity of the FED policy analysis, \( W_d(s) = 1 \) is taken for this section.

The uncertainty weight function is to be determined according to model uncertainty characteristics such that low and high frequency contribution of the uncer-
tainty should be reflected in this function. This example assumes that

\[ W_G(s) = \frac{s + 0.1}{s + 1} \]  \hspace{1cm} (4.22)

weighting function that reflects the 10% uncertainty at the low frequencies and 100% uncertainty for the high frequencies. These weighting functions are given in Fig.4.7.

![Figure 4.7](image)

**Figure 4.7.** The uncertainty \( W_G(s) \) and performance \( W_e(s) \) weighting functions.

Since the weighting functions are properly defined, the control system in Fig.4.6 can be converted into an LFT framework as shown in Fig.4.4. The overall framework is shown in Fig.4.8 with \( P(s) \) is given in Eqn.4.23.

\[
\begin{bmatrix}
  z(s) \\
  e(s) \\
  y(s)
\end{bmatrix} =
\begin{bmatrix}
  0 & 0 & G_N(s)W_G(s) \\
  W_e(s) & W_d(s)W_e(s) & G_N(s)W_e(s) \\
  -1 & -W_d(s) & -G_N(s)
\end{bmatrix}
\begin{bmatrix}
  w(s) \\
  d(s) \\
  u(s)
\end{bmatrix} \hspace{1cm} (4.23)
\]

The LFT structure in Eqn.4.23 is solved by using the technical steps in Section 4.2 and Matlab \( \mu \)-Toolbox is used during \( \mu \)-synthesis calculations to design a controller and related bounds for \( \mathbb{H}_\infty \) and \( \mu \) bounds are given in Fig.4.9. As seen from the figure, \( \mu \) value is always less or equal to \( \mathbb{H}_\infty \) bound. Moreover, the \( \mu \) value is
Figure 4.8. The macroeconomic stabilization problem in robust control framework.

found to be 0.4739 which is less than 1 and it implies the robust performance.

Figure 4.9. The $\mathcal{H}_\infty$ and $\mu$ bounds.

It should be noted that when the overall system is simulated for reference inputs, all weight functions are removed and the loop is closed, as shown in Fig.4.10.

The closed loop system is simulated assuming that the historical inflation is the reference input($R[z]$) and the actual estimate of the inflation is the output variable($Y[z]$) and the FED fund rate change is the control input($U[z]$). $\frac{Y[z]}{R[z]}$ and $\frac{U[z]}{R[z]}$ transfer functions are simulated and given in Fig.4.11, respectively. It is obvi-
Figure 4.10. The macroeconomic stabilization closed loop system.

Since the reference signal was historical inflation rates, it is expected to be the same at the output if the FED policy changes were to be optimal with respect to uncertainty and performance objectives. To show this case, the $W_e(s)$ weighting function is adjusted to achieve perfect tracking and the FED fund rate changes are calculated to compare the optimal and actual historical FED fund rate changes. In this optimal case, $W_e(s) = 990 \frac{s+1}{s+1}$ is chosen and the $\mu$ value is found to be 0.9972. The corresponding inflation tracking and FED fund rate changes are given in Fig.4.12. As seen from the figure, actual FED fund rate changes are much
more smoother and smaller in magnitude than the optimally suggested values, implying that the very cautious central bank policy changes. Summarily, under the case of model uncertainty, FED policy tends to change more cautiously and less reactionary, as most economists believe.

Figure 4.12. The optimal inflation tracking and FED fund rate changes.

Once the $\mu$-controller is designed, this optimization is applied in model (in)validation framework such that as long as the model is not invalidated, the optimization rule is still valid and should be continued. However, as soon as the model is invalidated, this means that the uncertainty bound is too big with respect to the presumed bound during optimization and new model identification-(in)validation should be done to determine a new controller. One obstacle is that the model (in)validation for structured uncertainty yields Bi-Affine Matrix Inequality(BMI), which is NP-hard to solve[26]. Therefore, the model (in)validation should be done externally. Since the economic dynamics inevitably changes depending on policies, perceptions, etc., occasional recalculation of a new model dynamic and a new controller is inevitable in this framework. The proposed Linear Parameter Varying(LPV), studied in Chapter 5, setup overcomes this problem.
4.4 Conclusions and Future Research

The main result of this chapter is that LFT approach in a $\mu$-controller framework can be used to overcome uncertainty issues in macroeconomic stabilization. $\mu$-controller framework and Bayesian analysis results, gain scheduling results and rule-based policy results can be compared. Also, the different stability and performance measures such as $\ell_1 - \mathcal{H}_\infty$, a multi-objective problem, can be used to better optimize macroeconomic stabilization problems.

The semi-blind model identification can be extended for multi-input multi-output systems, where the FED funds rate, the government spending and the overall tax ratio are the inputs, inflation, Gross-Domestic products, unemployment, etc. are the outputs. Afterwards, this system can be used in $\mu$-controller and model invalidation frameworks together to better represent the overall economic dynamic.

Although the coefficients are fixed in $G[z]$, it may be assumed that these coefficients are varying in some range. By using LFT approach, these deviations can be handled as an uncertainty and the parametric uncertainty structure can be obtained\textsuperscript{9}. This range could be a one-standard deviation, two-standard deviation so that the accuracy of the results can be given statistical interpretation as well.

Onatski and Stock\textsuperscript{16} states several issues for future research; solution for a policy maker discounted loss function in structured uncertainty cases\textsuperscript{10}, improving uncertainty structure by working on maximum uncertainty bounds\textsuperscript{11} and computing the optimal rule within the set of rules with bounded risk. Also, Onatski assumes that the policy maker has a system uncertainty but private agents know the model. Also, Hansen and Sargent\textsuperscript{14} assume that both the government and the private agents share the same uncertainty, shown by $\theta$, for the model in their Ramsey problem solutions and presents different uncertainty descriptions as a generalization of their original solution. They find $\theta$ by using statistical methods and work on 2-$\theta$ model in which the government and the private sector have different uncertainty bounds.

\textsuperscript{9}The parametric uncertainty approach is different than model uncertainty where all sort of uncertainty is lumped into an uncertainty block while each parameter uncertainty is handled separately in the parametric uncertainty.

\textsuperscript{10}They are already working on this.

\textsuperscript{11}They find that if $r = r_{max}$ is used, the policy could be very conservative.
In this chapter, semi-blind linear parameter varying (LPV) system identification, model (in)validation and a controller design of a general macroeconomic stabilization problem are explained.

5.1 Semi-blind Linear Parameter Varying Identification

Control-oriented LPV identification has recently been one of the main research areas due to developments on LPV optimization techniques. There exist some LPV identification approaches to use rich theoretical LPV optimization results, but this research will focus on LMI approach control-oriented LPV identification[49] and will further develop the identification by including initial conditions. The main idea is that the system past inputs can be used to represent the initial conditions through a known Hankel operator, as explained in Chapter 2, for the nonparametric dynamics and reachable set concept[66] can be used to sample initial conditions for the LPV dynamics.

The block diagram in Fig.5.1 can be used to show LPV identification with initial conditions. In the Fig.5.1, $u^+$ and $u^-$ denote system present and past input vectors, respectively, $y$ denotes the system output vector corrupted by noise $\eta \in \mathcal{N}$,
where \( \mathcal{N} = \{ \eta \in \mathbb{R}^n : |\eta| \leq \epsilon \} \), \( x^{ic} \) denotes the sampled initial states from the reachable set of the parametric part of the system, the linear parameter varying block \( \varphi \)

\[
\varphi = \begin{bmatrix}
\phi_1 I_{r_1} & & \\
& \ddots & \\
& & \phi_s I_{r_s}
\end{bmatrix}.
\]  

(5.1)

where \( \phi = \{ \phi \}_{k=0}^\infty \) denote varying parameters, is in diagonal form, \( G_p \) is the parametric transfer function set, \( G_{np} \in \mathcal{BH}_{\infty, \rho}(K) \) is the nonparametric part of the system. Then, for the system transfer function,

\[
G = \mathcal{F}_u(G_p, \varphi) + G_{np}
\]  

(5.2)

can be written by using the upper Linear Fractional transformation. The first term in Eqn.5.2 can be interpreted as the parameter varying part of the whole system. As done before[49], the parameter varying part of the system is assumed to have the expansion

\[
\mathcal{F}_u(G_p, \varphi) = \sum_{j=1}^{N_p} p_j \mathcal{F}_u(G_j, \varphi)
\]  

(5.3)

where \( p_j \) denotes unknown scalars and \( G_j \) denotes the known system matrices

\footnote{This is a very usual assumption for the nonparametric part of a system in robust identification.}
whose impulse responses are linearly independent for all parameter trajectories.

**Definition 5.1.1 (Consistency Set).** The set of all possible models coherent with the a-priori assumptions \((S, N)\), that could generate the a-posteriori experimental data is called the consistency set and is defined as

\[
T(y, \varphi) = \{ G \in T : y - T_N^G u \in N \}
\]

Linear parameter varying identification problem achieves its goal by establishing consistency and finding a nominal model. As done in Chapter 2, this could be presented as a Carathéodory-Fejér interpolation problem, which turns into a feasibility problem.

**Theorem 5.1.1 (LPV Consistency).** For given \((u^+, y, \phi)\), the consistency set has at least one element, i.e., \(T(y, \varphi) \neq \emptyset\), if and only if there exist two vector sequences \(p = [p_1 \, p_2 \, \ldots \, p_{Np}]^T\) and \(h = [h_0 \, h_1 \, \ldots \, h_{N-1}]^T\) satisfying

\[
M_R(h) \geq 0 \tag{5.4}
\]

\[
y = \sum_{j=1}^{Np} P_j \left[ T_u^+(G_j, \varphi) u^+ + T_u^-(G_j, \varphi) x^{ic} \right] + T_{\epsilon_{np}}^+ u^+ + T_{\epsilon_{np}}^- u^- + \eta \tag{5.5}
\]

where

\[
R = \begin{bmatrix}
1 \\
\rho \\
\rho^2 \\
\vdots \\
\rho^{N-1}
\end{bmatrix}, \quad F = \begin{bmatrix}
\epsilon_0 \\
0 \\
0 \ 0 \ h_0 \ h_1 \ h_2 \ h_3 \\
\vdots \\
0 \ 0 \ 0 \ 0 \ h_0 \ h_1 \ h_2 \ h_3 \ h_4 \ \cdots \ h_N \ z^N
\end{bmatrix}, \quad \eta \leq \epsilon \tag{5.6}
\]

\[
G_{\epsilon_{np}} = h_0 + h_1 z + h_2 z^2 + \ldots + h_N z^N + \ldots \tag{5.7}
\]
\[ T_{\mathsf{G}_{np}}^+ = \begin{bmatrix} h_0 & 0 & 0 & 0 & \ldots \\ h_1 & h_0 & 0 & 0 & \ldots \\ h_2 & h_1 & h_0 & 0 & \ldots \\ h_3 & h_2 & h_1 & h_0 & 0 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \end{bmatrix} \quad T_{\mathsf{G}_{np}}^- = \begin{bmatrix} h_1 & h_2 & h_3 & h_4 & h_5 & \ldots \\ h_2 & h_3 & h_4 & h_5 & h_6 & \ldots \\ h_3 & h_4 & h_5 & h_6 & h_7 & \ldots \\ h_4 & h_5 & h_6 & h_7 & h_8 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \end{bmatrix} \] (5.8)

\( T_{\mathsf{G}_{np}}^+ \) and \( T_{\mathsf{G}_{np}}^- \) are the pairs of Toeplitz and Hankel operators for the associated inputs \( u^+ \) and \( u^- \), respectively. For the current input \( u_{\mathsf{Nt} \times \mathsf{1}}^+ \), the corresponding Toeplitz operators can be found by picking the upper left \( (\mathsf{Nt} \times \mathsf{Nt}) \) matrix from \( T_{\mathsf{G}_{np}}^+ \) matrix. On the other hand, the Hankel operators can be found by taking the current input length \( (\mathsf{Nt}) \) as its row number and the order of the \( \mathsf{G}_{np} \) systems as its column number and picking the upper left matrix of this size from \( T_{\mathsf{G}_{np}}^- \) matrix. Since \( \mathsf{G}_{np} \) can have any order, the corresponding Hankel operator \( T_{\mathsf{G}_{np}}^- \) will have the number of columns, accordingly.

\( T_{\mathcal{F}_u(\mathsf{G}_{j,\varphi})}^+ u^+ \) term comes from parameter varying system simulation over the parameter trajectory. \( T_{\mathcal{F}_u(\mathsf{G}_{j,\varphi})}^- u^{ic} \) shows initial condition response that can be found by simulating the parameter varying system for some initial states which are sampled from the system reachable states.

**Proof:** Application of Carathéodory-Fejér interpolation theory results with regard to the existence of an operator \( \mathsf{G}_{np} \) is equal to \( \mathbf{R}^{-2} - \frac{F^TR^{-2}F}{K^2} \geq 0 \) matrix inequality condition[61]. Multiplying both sides of the inequality by \( K \) and using Schur’s compliment to this inequality yields Eqn.5.4. Eqn.5.5 simply comes from the consistency condition.

Nonparametric nominal model can be found by taking \( \mathsf{G}_{np} = C(zI - A)^{-1}B + D \), where \( h_0 = D, \ h_k = C.A^{k-1}.B \) for \( k \geq 1 \). Consequently, the system model is found to be \( \mathsf{G} = \sum_{j=1}^{\mathsf{Np}} p_j \mathcal{F}_u(\mathsf{G}_j, \varphi) + \mathsf{G}_{np} \).

### 5.2 Macroeconomic System LPV Identification

Knowing that the Federal Reserve Bank (FED) as one of the policy makers along with government, the changing nature of the inflation-FED funds rates dynamics
can be better represented by using LPV identification framework, utilizing varying economic indicators as parameters for different operating conditions. However, the question of accurate indicators of the overall economy, whether short-term real rates, leading economic indicators, National Association of Purchasing Managers (NAPM) index, company earning reports, output gap, etc., i.e., lagging, current and leading indicators, is still an open question, is completely arguable and economists dispute on these indicators and dynamic models. Yet, many different indicators are currently used by financial institutions for policy analysis and forecasting. However, it is essential to state that the number of indicators used should be low, 2 or 3, for efficient LPV controller design before getting into tractability issues of LPV framework. Different indicator groups can be taken in separate designs and the results can be compared to overcome this problem.

This study uses the output gap as the varying economic indicator in LPV identification and this assumption is coherent with the Phillips curve, i.e., inflation output gap dynamics, and IS curve, i.e., output gap to real-interest rate dynamics. By following the semi-blind LPV identification method, in Fig. 5.1 and explained

2 Lagging, current and leading indicators show the past, current and expected near future states of the economy, respectively. For example, unemployment numbers are lagging indicators due to the fact that they reflect slowing economy after it occurs. On the other hand, NAPM index as well as working hours are key indicators of the future state of the economy.
in Section 5.1 and using 14 data points (1961.1 and 1964.2 quarters) as shown in Fig.5.2, $u^+$ is taken as the FED funds rate, $u^-$ is the past FED funds rate and corresponding initial state conditions are used in an inequality conditions shown in Section 2.3, $x^{ic}$ is the sampled initial states for the corresponding subsystems of the parametric part, $y$ is the annual percentage inflation rate, $\phi$ is the output gap which is also the varying parameter of the system. The transfer functions associated with the parameter varying part can be found by using the idea that the parameter varying system can be decomposed into linear, time-invariant systems for relatively constant varying parameter and the overall behavior of the parameter varying system can be obtained by interpolation or extrapolation. In other words, for a constant $\phi_1$ over a data range, a $g_1(n) \Leftrightarrow G_1[z]$ transfer function can be found to cover the maximum power region by using the data frequency analysis. Similarly, another $g_2(n) \Leftrightarrow G_2[z]$ can be found by using the power spectrum analysis for another data range for a constant $\phi_2$. The overall behavior of parameter varying system during the identification range can then be seen as an interpolation of these two subsystems, i.e., $g = a_1[ (\phi - \phi_2).g_1 + (\phi - \phi_1).g_2 ]$, where $a_1$ is the scaling factor. By using this approach, the average $\phi_1 = 6.9670$ and

$$G_1(z) = \frac{0.1438z^2 + 0.2876z + 0.1438}{z^2 - 0.005628z - 0.2273}$$

(5.9)

were determined for the first three points, and the average $\phi_2 = 9.6746$ and

$$G_2(z) = \frac{0.2976z^2 + 0.5953z + 0.2976}{z^2 - 0.09509z + 0.6069}$$

(5.10)

were found for the remaining 11 data points, for the parameter varying part transfer functions\(^3\). The parameter varying part of the system can be expressed as

$$g_{lpv} = a_1[ \underbrace{ (\phi_1 - \phi_2).g_1 + \phi_1 (g_1 - g_2) }_{K_1} + \phi_2 (g_1 - g_2) ]$$

(5.11)

which could be represented in linear fractional transformation form as

$$\mathcal{F}_u(g, \phi) = K_1 + \phi.K_2, \text{ where } g = \begin{bmatrix} 0 & K_2 \\ K_1 & 1 \end{bmatrix}$$

(5.12)

\(^3\)During the model calculations, all continuous to discrete and discrete to continuous transformations are done by using bilinear approximation.
To obtain \( x^{ic} \) sampled initial conditions, \( G_1(z) \) and \( G_2(z) \) are used to calculate the smallest ellipsoids that contain the associated system reachable sets by solving [66, 67]

\[
W - A.W.A^T = B.B^T
\]

(5.13)

where \( W \) is the Lyapunov function and \( A, B \) are the matrices associated with \( G(z) = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \), for these transfer functions. Then, the size of the smallest ellipsoid containing the corresponding system reachable set, determined by the size of \( W \), is used to obtain initial condition samples by following the technique described in [68]. Finally, these initial condition samples are applied to \( G_1(z) \) and \( G_2(z) \) systems and the LPV system initial condition response can be obtained by using \( g_{lpv} \) in Eqn.5.11.

The input-output Fourier transforms are studied and \( \rho = 1.1 \) is chosen to reflect the stability margin of the the nonparametric part of the system. Since actual inflation calculation is not an easy task, \( \eta = 0.2 \) is set for the noise bound, corrupting the true inflation. Since the average value of the inflation is 1.1781 during the identification period, noise bound corresponds to \( \approx 20\% \) error margin. The time varying parameter is shifted up by 10 to avoid numerical problems during differential equation solutions.

![Figure 5.3. The historical data and the LPV identification-forecasting inflation outputs.](image)

Matlab LMI toolbox was used to perform the LPV identification calculations,
explained in the theorem 5.1.1 in Section 5.1, yielding $K = 0.2340$, $p_1 = 0.5412$, and an 11th order nonparametric system. By using the LPV system description $\mathbf{g} = p_1 \mathcal{F}_u(g, \phi) + g_{np}$, the historical data and appropriate samples for initial conditions, the LPV system identification and forecasting outputs are calculated and given in Fig.5.3, which clearly shows that the LPV identification and reasonably good forecasting are successfully achieved. When compared with the linear, time-invariant model response given in Fig.2.4, the LPV model has a maximum error of 0.8988 and the linear, time-invariant model has approximately 1.2. Since the LPV system used only one time-varying parameter, two parametric transfer functions, small data range during the identification, degraded forecasting performance for the longer data range suggests that more time-varying parameters, more accurate parametric transfer functions and longer data range should be used to improve the forecasting accuracy.

5.3 Linear Parameter Varying Model (In)validation

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5.4.png}
\caption{The semi-blind LPV Model (In)validation setup.}
\end{figure}

The LPV model (in)validation setup can be illustrated by using the Fig.5.4, where the convex relaxation principle is used to avoid nonlinear matrix inequalities\textsuperscript{4}. For any $(u^-, u^+, \tilde{\eta}, \delta)$ satisfying

$$
\mathbf{y} = (1 + \Delta) \left( p_1 \left[ T_{+}^{+} F_{+} (g, \phi) u^+ + T_{+}^{-} F_{-} (g, \phi) x^{ic} \right] + T_{+}^{+} g_{np} u^+ + T_{-}^{-} g_{np} u^- + \tilde{\eta} \right) \quad (5.14)
$$

\textsuperscript{4}Convex relaxation, i.e., moving the noise input before the uncertainty block, is explained in Section 3.3 on page 32. Moreover, by using the steps in Section 3.2, it can be easily shown that the setup before convex relaxation yields nonlinear matrix inequalities during semi-blind validation.
where \(|\Delta|_\infty \leq \delta, |\tilde{\eta}| \leq \tilde{\epsilon} = \frac{\epsilon}{1+\delta}\) and \(T^+_{(q)}\) and \(T^-_{(q)}\) are the appropriate operators to represent current and past input responses, respectively, then the \((u^-, u^+, \eta, \delta)\) terms with \(\eta \doteq (1 + \Delta)\tilde{\eta}\) satisfies the setup before convex relaxation. Therefore, the model invalidation process can be achieved by searching for a solution of the problem setup, shown in Fig.5.4. Since the uncertainty bound is expected to be small such as \(|\Delta|_\infty \ll 1\), it is assumed that the original noise bound can be used after the modification and the new setup is not too conservative.

**Problem 5.3.1 (Semi-blind LPV Model (in)validation).** For given time domain data vectors \(u^+ = \{u_j\}_{j=0}^{N-1}, \phi = \{\phi_j\}_{j=0}^{N-1}, y = \{y_j\}_{j=0}^{N-1}\), the LPV model \(g(\phi) = p_1 \mathcal{F}_u(g, \varphi) + g_{np}\), the a-priori sets \(\tilde{\eta} \in \tilde{N}, \tilde{N} \doteq \{\tilde{\eta}; \tilde{\eta} \leq \epsilon\}\) and \(\Delta \in \Delta\), where \(\Delta = B\mathcal{H}_{\infty, \rho}(K)\) determine whether the set

\[
S(y, \phi, g) = \{(\tilde{\eta}, \Delta) : \tilde{\eta} \in \tilde{N}, \delta \in \Delta, y = T_{g(\phi)}u + \tilde{\eta}\}
\]

is nonempty, i.e., a-priori and a-posteriori information is consistent.

**Theorem 5.3.1 (LPV Model (in)validation).** For given \((u^+, \phi, y)\) time domain data vectors, the \(g(\phi)\) LPV system model is not invalidated by these vectors if and only if there exist \(\tilde{\eta} = [\tilde{\eta}_0 \tilde{\eta}_1 \ldots \tilde{\eta}_{N-1}]^T\) and \(\omega = [\omega_0 \omega_1 \ldots \omega_{N-1}]^T\) vectors such that

\[
M^k(\omega) \doteq \begin{bmatrix} X^k(\omega) & (\omega^k)^T \\ (\omega^k) & -(\delta^2 - 1)^{-1}\mathbf{I} \end{bmatrix} \leq 0 \quad k = 1, 2, \ldots, N
\]

\((N+1)\) LMIs hold, where

\[
\begin{align*}
X^k(\omega) & \doteq (y^k)^T y^k - (y^k)^T \omega^k - (\omega^k)^T y^k \\
\omega^k & \doteq p_1 \left[ T^+_{\mathcal{F}_u(g, \varphi)} u^+ + T^-_{\mathcal{F}_u(g, \varphi)} x_{ic} \right] + T^+_{g_{np}} u^+ + T^-_{g_{np}} u^- + \tilde{\eta} \\
\tilde{\eta} & \doteq \omega - z \leq \epsilon
\end{align*}
\]

**Proof :**

The existence of an uncertainty block in Fig.5.4 is equivalent to[62],

\[
(\xi^k)^T \xi^k \leq \delta^2 (\omega^k)^T \omega^k \quad k = 1, 2, \ldots, N
\]
By using the setup in Fig.5.4, \( \xi^k = y^k - \omega^k \) can be written and be substituted in Eqn.5.16 yielding

\[
(y^k - \omega^k)^T (y^k - \omega^k) \leq \delta^2 (\omega^k)^T \omega^k \quad k = 1, 2, \ldots, N \quad (5.17)
\]

Rearranging the terms,

\[
(y^k)^T y^k - (y^k)^T \omega^k - (\omega^k)^T y^k \leq (\delta^2 - 1)(\omega^k)^T \omega^k \quad k = 1, 2, \ldots, N \quad (5.18)
\]

can be written. By using the Schur’s complement,

\[
M^k(\omega) \doteq \begin{bmatrix}
(y^k)^T y^k - (y^k)^T \omega^k - (\omega^k)^T y^k & (\omega^k)^T \\
\omega^k & -(\delta^2 - 1)^{-1}
\end{bmatrix} \leq 0 \quad (5.19)
\]

where \( \omega^k \doteq p_1 \left[ T_{\mathcal{F}_u(G_j,\phi)} u^+ + T_{\mathcal{F}_u(G_j,\phi)} x^{ie} \right] + T_{G_{np}} u^+ + T_{G_{np}} u^- + \tilde{\eta} \), can be obtained. By comparing the each term in Theorem 5.3.1 and in Eqn.5.19, the proof is completed.

### 5.4 Macroeconomic System LPV Model (In)validation

Even if the identification results are very good, the LPV system should be tested by using totally new data to measure its forecasting power. Robust control theory model (in)validation concept quantifies the forecasting power of a model.

Matlab LMI toolbox was used to perform the semi-blind LPV model validation calculations, explained in the theorem 5.3.1 in Section 5.3, for the identification data range and \( \delta = 0.0001 \) was found, which is expected for the identification data range. The initial conditions were the same for the nonparametric and time varying systems, as explained in Section 5.2. For the 1-25 data points, the LPV model is still validated with the \( \delta = 0.0886 \), which is very reasonable uncertainty bound with respect to robust control approach. However, degraded model (in)validation results for the longer data range suggest that more time-varying parameters, more accurate parametric transfer functions and longer data range should be used to improve the validation results.
5.5 Linear Parameter Varying Controller Design

LPV controller framework offers great advantages to control nonlinear or time-varying systems. For the affine parameter dependent systems, which could be a time varying system or a linearized nonlinear system, gain-scheduled $H_\infty$ framework\(^5\) is utilized to design a controller for each operating point, represented by the linear varying parameter(s) which are assumed to be measured in real-time, and these LPV controllers are scheduled by the corresponding operating points. An LPV controller achieves better performance than a single linear time-invariant controller that is designed for the whole parameter variation range. Another important benefit of the LPV framework is that part of the states can be treated as linearly varying parameter provided that these states can be measured in real-time.

Consider the LPV model family,

$$
\begin{bmatrix}
\dot{x}(t) \\
z(t) \\
y(t)
\end{bmatrix} =
\begin{bmatrix}
A(\phi(t)) & B_1(\phi(t)) & B_2 \\
C_1(\phi(t)) & D_{11}(\phi(t)) & D_{12} \\
C_2 & D_{21} & D_{22}
\end{bmatrix}
\begin{bmatrix}
x(t) \\
\omega(t) \\
u(t)
\end{bmatrix}
$$

(5.20)

where $\phi(t) = [\phi_1(t) \ \phi_2(t) \ \ldots \ \phi_n(t)]$ is the vector of time varying parameter trajectories satisfying $\phi(t) \in \mathcal{P}$ for all $t$ and $\nu_i(\phi(t)) \leq \dot{\phi}_i(t) \leq \nu_i(\phi(t))$ holds for each $i$ and all $t$ and $A(\phi)$, $B_i(\phi)$, $C_i(\phi)$ and $D_{ij}(\phi)$ are affine parameter dependent system matrices. While $x(t)$ denotes the states of the system, $\omega-z$ pair represents the performance input and outputs and $u-y$ pair denotes the LPV controller input and outputs.

When the parameter vector $\phi(t)$ varies in a box of $\mathbb{R}^n$ with corners $(\pi_i)_{i=1}^N$ where $N = 2^n$, by using convex decomposition,

$$
\phi(t) = \alpha_1 \phi_1 + \alpha_2 \phi_2 + \ldots + \alpha_N \phi_N
$$

(5.21)

where $\alpha_i > 0$ and $\sum_{i=1}^N \alpha_i = 1$, can be written. It should be noted that as the $\phi(t)$ vector varies, the system matrix $S(\phi)$ ranges in a matrix polytope with $S(\pi)$

\(^5\)The fundamental part of this section is interpreted from [54] and [63].
vertices such that

\[
S(\phi) = \alpha_1 S(\pi_1) + \alpha_2 S(\pi_2) + \ldots + \alpha_N S(\pi_N)
\] (5.22)

can be obtained. For a general LPV controller state-space equations,

\[
\begin{bmatrix}
\dot{x}_K(t) \\
u(t)
\end{bmatrix} =
\begin{bmatrix}
A_K(\phi(t)) & B_K(\phi(t)) \\
C_K(\phi(t)) & D_K(\phi(t))
\end{bmatrix}
\begin{bmatrix}
x_K(t) \\
y(t)
\end{bmatrix}
\] (5.23)

can be written and different \(K(\pi)\) controllers can be designed to satisfy the performance objectives for the vertices of \(S(\phi)\). By using the convex decomposition of these controllers, the actual LPV controller for different time varying parameters, i.e., operating points, can be expressed as

\[
K(\phi) = \alpha_1 K(\pi_1) + \alpha_2 K(\pi_2) + \ldots + \alpha_N K(\pi_N)
\] (5.24)

and the smooth scheduling of the LPV controller can be achieved. It should be noted that the controller depends only on \(\phi\) when state feedback is used whereas it depends on \(\phi\) and \(\dot{\phi}\) when output feedback is used, though the \(\dot{\phi}\) dependence can be removed in special circumstances[54]. The linear varying parameter L2-Induced gain theorem, theorem D.2.1 on page 90, is the starting point for LPV controller synthesis when state or output feedback is used. Although it has very appealing properties of LPV analysis, there are some limitations such that the compact inequality given in Eqn.D.7 represents \(2^m\) inequalities coming from partial derivatives and it must hold for \(\forall \phi \in \mathcal{P}\) and all possible matrix functions of a parameter vector \(\phi\) should be used when searching for \(X\).

Since the gain-scheduled \(H_\infty\) approach, one way to implement LPV controllers, is explained above, the basic LPV controller synthesis involving general block diagram can be introduced by using the Fig.5.5 and the LPV tracking problem can be stated as given in Problem 5.5.1,

where \(r\) denotes the reference input, \(e\) denotes the tracking error, \(u\) denotes the plant input, \(d\) denotes the disturbance, \(y\) denotes the output, \(K(\phi)\) denotes an LPV controller, \(G_N(\phi)\) denotes the LPV nominal model, \(\Delta_G\) denotes the norm-bounded
LPV system uncertainty.

**Problem 5.5.1 (LPV Tracking Problem).** Does there exist a parameter dependent controller $K(\phi)$ to satisfy the performance objectives given below;

- Minimum $L_2$-to-$L_2$ worst-case gain from disturbance to the output,
- Minimum $L_2$-to-$L_2$ worst-case gain from $\xi$ to $\omega$ to ensure robustness against the model uncertainty

for all allowable parameter trajectories, for the system shown in Fig.5.5.

During a controller design, these performance objectives are converted to rms gain constraints such that

$$
\frac{Y(s)}{D(s)} = \frac{1}{1 + K(\phi).G_N(\phi)} \equiv \frac{E(s)}{R(s)} \equiv S(s)
$$

(5.25)

where $S(s)$ is the sensitivity function, can be written for the disturbance rejection whose rms gain constraint\(^6\)

$$
||W_1(s).S(s)||_{\infty} < 1
$$

(5.26)

where $W_1(s)$ is the weighting function, can be obtained. By using the small-gain theorem, an rms gain constraint for the robustness against the model uncertainty

\(^6||S(s)||_{2\rightarrow 2} = \sup_{||d||_2 \leq 1} \frac{||S(s)d||_2}{||d||_2}\) represents the worst case, i.e., induced $L_2$ norm, which is equal to $||S(s)||_{\infty}$.
is equal to
\[ ||W_2(s)T(s)||_\infty < 1 \quad (5.27) \]
where \( W_2(s) \) is the uncertainty weighting filter and \( T(s) = \frac{K(\phi).G_N(\phi)}{1+K(\phi).G_N(\phi)} \) is the complementary sensitivity function. Finally, these constraints are combined together for the tractability of \( H_\infty \)-framework yielding
\[
\begin{bmatrix}
W_1(s).S(s) \\
W_2(s).T(s)
\end{bmatrix}
\leq 1
\quad (5.28)
\]

Once the appropriate transfer functions are determined, \( H_\infty \) calculations can be carried out provided that the linear fractional representation of the augmented plant, uncertainty and the controller is established. Since the controller and the uncertainty is not known initially, except the uncertainty bound, these blocks are removed and their inputs-outputs including performance inputs-outputs are used to determine the augmented plant. When the uncertainty and the controller are removed and the weighting functions are added properly, by using the Fig.5.5, either \( r \) or \( d \), \( u \) and \( \xi \) as the inputs and \( e \), \( \omega \) and \( y \) as the outputs, the augmented plant can be calculated and the corresponding LFT structure can be given, as shown in Fig.5.6, to perform the \( H_\infty \) calculations.

\[ \text{Figure 5.6. The Linear Parameter Varying Synthesis Framework.} \]
As seen in the Fig.5.6, $P(\phi)$ is the nominal model and $P_{aug}(\phi)$ is the augmented plant including the weighting filters, which is not part of the actual physical system. It should be noted that once an LPV controller designed, only $P(\phi)$, the actual system, is used to test the controller performance.

The proposed controller design, given in Fig.5.6, utilizes the output feedback and the LPV controller synthesis follows applying the theorem D.2.1 for the closed loop system matrices, i.e., the system and the controller matrices in output feedback form. The existence of a controller for the affine parameter varying system output feedback $|\mathcal{F}(P_{aug}(\phi), K(\phi))|_\infty \leq \gamma$ problem depends on[64, 65, 63, 54] the existence of two symmetric $X$ and $Y$ matrices such that

\[
\begin{bmatrix}
N_{12} & 0 \\
0 & I
\end{bmatrix}^T
\begin{bmatrix}
A_i X + X A_i^T & X C_i^T & B_{1i} \\
C_{1i} X & -\gamma I & D_{11i} \\
B_{1i}^T & D_{11i}^T & -\gamma I
\end{bmatrix}
\begin{bmatrix}
N_{12} & 0 \\
0 & I
\end{bmatrix} \prec 0, \quad i = 1, 2, ..., N
\tag{5.29}
\]

\[
\begin{bmatrix}
N_{21} & 0 \\
0 & I
\end{bmatrix}^T
\begin{bmatrix}
A_i^T Y + Y A_i & Y B_{1i} & C_{1i}^T \\
B_{1i}^T Y & -\gamma I & D_{11i}^T \\
C_{1i} & D_{11i} & -\gamma I
\end{bmatrix}
\begin{bmatrix}
N_{21} & 0 \\
0 & I
\end{bmatrix} \prec 0, \quad i = 1, 2, ..., N
\tag{5.30}
\]

\[
\begin{bmatrix}
X & I \\
I & Y
\end{bmatrix} \succeq 0
\tag{5.31}
\]

where

\[
\begin{bmatrix}
A_i & B_{1i} \\
C_{1i} & D_{11i}
\end{bmatrix} \doteq \begin{bmatrix}
A(\pi_i) & B_1(\pi_i) \\
C_1(\pi_i) & D_{11}(\pi_i)
\end{bmatrix}
\]

and

\[
[B_2^T \ D_{12}^T] N_{12} = 0 \quad \text{and} \quad [C_2 \ D_{21}] N_{21} = 0
\]

i.e., $N_{12}$ is the basis of the null-space of $[B_2^T \ D_{12}^T]$ and $N_{21}$ is the basis of the null-space of $[C_2 \ D_{21}]$. 

5.6 Macroeconomic Stabilization by LPV Controllers

The proposed LPV controller synthesis framework can be used to stabilize macroeconomic variables. Assuming the Federal Reserve Bank (FED) as the policy maker with the control input of the FED funds rates, macroeconomic stabilization by inflation targeting can be dealt by LPV optimization framework. Being compatible with the generally accepted aggregate supply and aggregate demand equations, i.e., Phillips and IS equations, the gross-domestic product output gap is taken to be the linearly varying parameter to represent the different operating points of the FED funds rate-inflation dynamics. It is extremely important to say that there should be 2 or 3 linearly varying parameters for efficient LPV optimization before getting into tractability issues. However, different economic indicator groups can be taken in separate optimization designs and the results can be compared to tackle this problem.

Following the Section 5.5, the macroeconomic stabilization problem can be proposed by using the Fig.5.7,

\[
\begin{align*}
\hat{e} & \rightarrow W_1(s) & e \rightarrow K(\phi) & u_{FED} \\
& \rightarrow G_N(\phi) & \omega & \rightarrow \Delta_G & \xi \\
& \rightarrow y_{act} & d & \rightarrow W_2(s) & \tilde{y}
\end{align*}
\]

**Figure 5.7.** Macroeconomic stabilization problem setup.

where \( r_{des} \) denotes the desired inflation trajectory, \( y_{act} \) denotes the actual inflation, \( u_{FED} \) denotes the FED funds rate, \( d \) denotes exogenous disturbance, \( G_N(\phi) \) is the nominal LPV system, \( \Delta_G \) is the multiplicative model uncertainty, \( W_1(s) \) and \( W_2(s) \) are the weighting functions for the corresponding variables.

The LPV controller framework is proposed as an inflation targeting problem for
25 data points, the historical inflation data between 1961.1 and 1967.2 quarters. For the nominal affine parameter-dependent model $G_N(\phi)$,

$$G_N(\phi) = p_1 \mathcal{F}_u(G, \varphi) + G_{np}$$

where $p_1$, $\mathcal{F}_u(G, \varphi)$ and $G_{np}$ are obtained during the semi-blind LPV model identification section, Section 5.2. $\|\Delta_G\|_{\infty} = 0.0886$ was used for the multiplicative uncertainty bound, which was found during the model invalidation Section 5.4. With respect to the weighting functions,

$$W_1(s) = \frac{s + 1}{\frac{s}{10} + 1} \quad W_2(s) = 0.1 \frac{s + 1}{\frac{s}{10} + 1}$$

are used so that the $e$ error signal is weighted by a low-pass filter and the $y$ output signal is weighted by a high-pass filter, as shown in Fig.5.6.

Once all the terms are identified in Fig.5.6, Matlab LMI toolbox is used to calculate a controller, actually by solving the Linear Matrix Inequalities given in Eqns.5.29, 5.30, 5.31. However, since the semi-blind LPV identification didn’t force $B_2$, $C_2$, $D_{12}$ and $D_{21}$ independent of the linearly varying parameter, these matrices initially depended on the linear varying parameter. As explained in the LMI toolbox, this restriction has been overcome by using low-pass filters on $u$ input and $y$ output paths. The designed controller is also parameter varying in nature and consisted 10 states. The LPV controller and the LFT structure not including $W_1(s)$ and $W_2(s)$ were used to generate the tracking framework, as shown in Fig.5.7, and the overall parameter-varying system time domain simulations were obtained by using historical inflation input as the reference input and using 'pdsimul' command in Matlab. The system output is given in Fig.5.8, where the top plot shows the historical data(solid) and LPV system output(dashed), while the bottom plot shows the LPV system error.

5.7 Conclusions and Future Research

The linear-parameter varying framework for semi-blind LPV model identification-validation and LPV controller synthesis are explained. Since the proposed LPV
framework incorporates the model changes indicated by a linear-varying parameter and accommodates it by scheduling the controller, then, the LPV framework is shown to be immune to Lucas critique to optimally stabilize the macroeconomy, represented by an affine parameter-varying system with a norm bounded multiplicative uncertainty.

Considering the Federal Reserve Bank (FED) and the government as the two policy makers, whose control inputs are the short-term interest rates, government spending and tax rates, respectively, of the overall economy, the general problem of destabilized economy, i.e., increasing inflation or unemployment, impact of sudden changes in disturbances, for different uncertainty structures, different number of quarters, can be dealt in in the LPV framework to study the effects of different scenarios.

LPV framework performance can be compared with Bayesian analysis results, rule-based policy results, switching regime systems, Ramsey problems [35], etc., and can be used in portfolio optimization, real-estate investment etc. Moreover, the question of useful indicators of the overall economy can be studied; if short-term real rates and leading economic indicators are taken to characterize the macroeconomic dynamics in LPV framework, do they yield the best results? Or, should some past variable averages and big company earning growth ratios be taken?

**Figure 5.8.** The LPV inflation targeting system outputs.
Since LPV framework can implicitly incorporate the forward-looking nature of the policy makers, individuals, etc., by including appropriate parameters, i.e., leading indicators or consumer confidence, LPV framework performance and performance of techniques using estimates of forward-looking variables can be studied.
Chapter 6

Conclusions

Convex relaxation frameworks for NP-hard problems of semi-blind $\ell_1$-identification, semi-blind model (in)validation and semi-blind linear parameter varying model identification, (in)validation and a controller design, the main contribution of the dissertation and an answer to Lucas critique have been successfully presented.

A $\mu$-controller has been designed for macroeconomic stabilization problems under model uncertainty.


MATLAB Programs

Due to the length of the Matlab files, any interested readers is respectfully asked to contact with the author at myilmaz@gandalf.ee.psu.edu.
Background Results on Robust Identification

The basic results on information based complexity (IBC) and robust identification concepts are summarized in this part. A comprehensive treatment of the subject as well as an extensive reference list can be found for instance in [28] or [26].

B.1 Worst Case Identification Error and Diameter of Information

A salient feature of robust identification is its ability to provide worst-case bounds on the identification error. Given an identification algorithm $\mathcal{A}$ mapping the a-priori and a-posteriori information to a candidate nominal model, its local error is defined as follows:

$$e(\mathcal{A}, y) = \sup_{g \in \mathcal{T}(y)} m[g, \mathcal{A}(y, S, N)]$$  \hspace{1cm} (B.1)

that is, the maximum distance between the identified model and any other system in the consistency set $\mathcal{T}(y)$. Note that this error is related to the outcome of a specific experiment $y$. A global error can be defined by considering the worst-case error over the set of all possible outcomes:

**Definition B.1.1.** The worst case global error of a given algorithm $\mathcal{A}(S, N, y)$ is
given by:

\[ e(\mathcal{A}) = \sup_{y \in Y} e(\mathcal{A}, y) \]  \hspace{1cm} (B.2)

where \( Y \) is the set of all possible data, consistent with sets \( S \) and \( \mathcal{N} \).

Next we briefly review how to obtain mathematically tractable bounds for these errors. Recall that the set \( \mathcal{T}(y) \subset S \) is the smallest set of models that are indistinguishable from the view point of the input information. Therefore, roughly speaking, its size gives lower and upper bounds on the identification error defined above. In order to formalize these concepts and obtain computable bounds we need to introduce the following definitions:

**Definition B.1.2.** The radius and diameter of a subset \( \mathcal{A} \) of a metric space \( (\mathcal{X}, m) \) are

\[ r(\mathcal{A}) = \inf_{x \in \mathcal{X}} \sup_{a \in \mathcal{A}} m(x, a) \]  \hspace{1cm} (B.3)  
\[ d(\mathcal{A}) = \sup_{x, a \in \mathcal{A}} m(x, a) \]  \hspace{1cm} (B.4)

The radius can be interpreted as the maximum error, measured in the metric \( m(,) \), when considering the set \( \mathcal{A} \) as represented by a single “central” point (which might not belong to \( \mathcal{A} \)). The diameter is the maximum distance between any two points in the set. Based on these concepts of radius, we next quantify the “size” of the available information.

**Definition B.1.3.** The radius and diameter of information are defined as:

\[ R(I) \doteq \sup_{y \in Y} r[\mathcal{T}(y)] \]  \hspace{1cm} (B.5)  
\[ D(I) \doteq \sup_{y \in Y} d[\mathcal{T}(y)] \]  \hspace{1cm} (B.6)

where \( Y \) is the set of all possible data consistent with the sets \( S \) and \( \mathcal{N} \):

\[ Y \doteq \{ E(g, \eta) \mid g \in \mathcal{S}, \eta \in \mathcal{N} \} \]  \hspace{1cm} (B.7)

The following result gives worst–case bounds of the identification error based on these concepts:
Lemma B.1.1. The worst case identification error defined in (B.2) satisfies the following inequality:

$$e(A) \geq R(I) \geq \frac{1}{2} D(I)$$  \hspace{1cm} (B.8)

for any algorithm $A$. The following upper bound holds:

$$D(I) \geq e(A_I)$$  \hspace{1cm} (B.9)

for any interpolation algorithm $A_I$.

The bounds above are of theoretical importance such that $R(I)$ can be interpreted as an intrinsic error that cannot be decreased by any identification algorithm, unless extra information is added to the problem. Although these quantities are in general hard to compute, they lead to mathematically tractable problems in practically relevant cases.

Definition B.1.4. A set $A$ in a linear space $X$ is called symmetric if and only if there exists an element $c \in X$ such that for any $a \in X$ for which $c + a \in A$ then $c - a \in A$. The element $c$ is called the symmetry point of set $A$.

Lemma B.1.2. If the a-priori sets $S$ and $N$ are symmetric and convex with respect to 0, and the experiment operator $E(g, \eta)$ is linear with respect to both $g$ and $\eta$ then the diameter of information satisfies:

$$D(I) = \sup_{y \in Y} d[T(y)] = d[T(y_0)] , \quad y_0 = E(0, 0)$$  \hspace{1cm} (B.10)

Furthermore,

$$d[T(y_0)] = 2 \sup_{g \in T(y_0)} m(g, 0)$$  \hspace{1cm} (B.11)

The above result states that the experiment that yields the least amount of information is the one that results in a null outcome. Moreover, a bound on the worst case identification error is given by twice the maximum distance from any element in $T(y_0)$ to the center of symmetry of $S$. 
B.2 Results on Interpolatory Algorithms

The first lemma considers the problem of the existence of a bounded causal linear discrete-time invariant operator that maps two given sequences.

Lemma B.2.1 (Tangential Carathéodory-Fejér Interpolation, [26]). Given vector valued sequences $u_i \in \mathbb{R}^l$ and $v_i \in \mathbb{R}^m$, $i = 0, 1, \ldots, n$, there exists a matrix interpolant $L(z) \in \mathcal{B} \mathcal{H}_{\infty}^{m \times l}$ such that

\[ L(z) = L_0 + L_1 z + L_2 z^2 + \ldots + L_n z^n + \ldots \]

\[ T^n_L u_i = v_i \]

if and only if $(T^n_u)^T T^n_u - (T^n_v)^T T^n_v \geq 0$.

The next lemma gives a necessary and sufficient condition for two finite vector sequences to be related by a system in a family $\mathcal{S}$ of the form given in Eqn.B.12:

\[ \mathcal{S} = \{ S(z) = F_{np}(z) + F_p(z) \} \]

where linear systems $S(z)$ are described in terms of a nonparametric component $F_{np}(z) \in \mathcal{B} \mathcal{H}_{\infty,p}(K)$ and a parametric component $F_p(z)$.

Lemma B.2.2 (Existence of an operator). Given a scalar $K_T$, and two vector sequences $(u, y)$, there exists an operator $S \in \mathcal{S}$ such that $y = Su$ if and only if there exists vectors $f, p$ satisfying:

\[ M(f) \doteq \begin{bmatrix} K_T R^{-2} (T^N_f)^T \\ T^N_f & K_T R^2 \end{bmatrix} \geq 0, \]

\[ y = Pp + T^N_u f \]

where $P$ and $R$ are defined as in Lemma 2.2.2. Moreover, in this case all such operators $S$ can be parameterized in terms of a free parameter $Q(z) \in \mathcal{B} \mathcal{H}_{\infty}$. In particular, the choice $Q(z) = 0$ leads to the "central" model

\[ S_{central}(z) = H_o(z) + p^T G_p(z) \]
where an explicit state–space realization of \( H_o(z) \) is given by:

\[
H_o(z) = C_H (zI - A_H)^{-1} B_H + D_H
\]

with

\[
\begin{align*}
A_H & = \{ A - [C^T C_- + (A^T - I)]^{-1}C^T C_- (A - I) \}^{-1} \\
B_H & = [C^T C_- (A^T - A - I) - (A^T - I)A]^{-1}C^T \\
C_H & = KC_+ - KC_+ \{ A - [C^T C_- + (A^T - I)]^{-1}C^T C_- (A - I) \}^{-1} \\
D_H & = KC_+ \{ [C^T C_- + (A^T - I)]A - C^T C_- (A - I) \}^{-1} C^T,
\end{align*}
\]

and

\[
A = \begin{bmatrix} 0_{I_N \times N} & N+1 \end{bmatrix}, \quad C_- = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}, \quad C_+ = \frac{h^T}{K}.
\]

Proof. See [53].
Historical Data

Since the last copy of the historical data from corresponding websites, it has been noticed that the data given for 1996 dollars have been adjusted to data for 2000 dollars, resulting in slight changes.

**Real GDP chained 1996 dollars** data are taken from both http://www.stls.frb.org/fred/data/gdp/gdpc1 and http://www.marketvector.com/data/index.htm (series ID=GDPC96) pages and data between 1960.4-2003.2 periods were used.

**Real Potential GDP** chained 1996 dollars are taken from http://www.stls.frb.org/fred/data/gdp/gdppot and data between 1960.4-2003.2 periods were used.

**Federal Funds Rate** Monthly averages of daily figures(Percent) data are taken from http://www.federalreserve.gov/Releases/H15/data/m/fedfund.txt and http://www.marketvector.com/data/index.htm (Series ID=FEDFUNDS) and the monthly average data between March 1960- July, 2003 periods were used.

**Gross Domestic Product:Chain-type Price Index** data are taken from http://www.marketvector.com/data/data/GDPCTPI.htm and data between 1960.1-2003.2 periods were used.
D.1 Small Gain Theorem

Theorem D.1.1 (Small Gain Theorem). [54] For a given system \( G = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} \) and \( B_\Delta = \{ \Delta \in \mathbb{R}^{n_1 \times n_1} : \sigma(\Delta) \leq 1 \} \), the necessary and sufficient condition for

\[
\max_{\Delta \in B_\Delta} \sigma[F_U(G, \Delta)] < 1
\]

is that both \( \sigma(G) < 1 \) and \( \sigma(\Delta) \leq 1 \) assuming that \( \sigma(G_{11}) < 1 \).

Proof: The underlying proof borrows from the idea that if the max. \( L_2 \) induced gains are less than 1, than the loop gain of the block diagram in Fig.D.1 will be less than one and \( \det(I - G.\Delta) \neq 0 \) will always hold yielding a stable system. Assumptions \( \sigma(G) < 1 \) and \( \sigma(\Delta) \leq 1 \) imply that \( (I - G.\Delta) \) is invertible and thus
the loop is well-posed. Then, for arbitrary vectors,
\[
\begin{bmatrix}
  w \\
  e
\end{bmatrix} = M \begin{bmatrix}
  z \\
  d
\end{bmatrix} \quad \text{and} \quad z = \Delta w \quad (D.1)
\]
can be written. Using the \( \sigma(M) \leq \alpha < 1 \) and \( \sigma(\Delta) \leq 1 \) assumptions, it is obvious that the vectors in Eqn.D.1 should satisfy
\[
\left|\left| \begin{bmatrix}
  w \\
  e
\end{bmatrix} \right|\right|^2 \leq \alpha^2 \left|\left| \begin{bmatrix}
  z \\
  d
\end{bmatrix} \right|\right|^2 \quad \text{and} \quad \left|\left| \begin{bmatrix}
  z \\
  d
\end{bmatrix} \right|\right|^2 \leq \left|\left| \begin{bmatrix}
  w \\
  e
\end{bmatrix} \right|\right|^2 \quad (D.2)
\]
Noting that \( e = F_U(G, \Delta)d \) and combining the inequalities\(^1\) in Eqn.D.2
\[
\left|\left| e \right|\right|^2 = \left|\left| F_U(G, \Delta)d \right|\right|^2 \leq \alpha^2 \left|\left| d \right|\right|^2 \quad (D.3)
\]
can be written. Since \( d \) is a nonzero arbitrary vector,
\[
\sigma[F_U(G, \Delta)] \leq \alpha < 1 \quad (D.4)
\]
is obtained. Since this result holds for all \( \Delta \in B_\Delta \),
\[
\max_{\Delta \in B_\Delta} \sigma[F_U(G, \Delta)] < 1 \quad (D.5)
\]
holds and the proof is completed.

A sufficient condition\(^2\) for the inequality given in Eqn.D.5 is that there exists a positive definite, hermitian, complex matrix \( D \in \mathbb{C}^{n_1 \times n_1} \), where \( n_1 \) is the size of the \( \Delta \) block, such that
\[
\sigma \left( \begin{bmatrix}
  D & 0 \\
  0 & I_{n_2}
\end{bmatrix} G \begin{bmatrix}
  D^{-1} & 0 \\
  0 & I_{n_2}
\end{bmatrix} \right) < 1 \quad (D.6)
\]
\(^1\)With regard to vector norms, \( \left|\left| \begin{bmatrix}
  w \\
  e
\end{bmatrix} \right|\right|^2 = \begin{bmatrix}
  w^T \\
  e^T
\end{bmatrix} \begin{bmatrix}
  w \\
  e
\end{bmatrix} = \| w \|^2 + \| e \|^2 \) holds.
\(^2\)This sufficient and necessary condition in Eqn.D.6 is related to the Bounded Real Lemma and Kalman-Yakubovic-Popov(KYP) Lemma. KYP Lemma establishes equivalent inequality conditions between frequency domain and state-space in terms of a Riccati equation or an LMI equation.
D.2 LPV $L_2$-Induced Gain Theorem

Theorem D.2.1 (LPV $L_2$-Induced Gain Theorem). \[54] Any LPV system $S(\phi)$ is internally exponentially stable and $||S(\phi)||_{2-2} < 1$ for all parameter trajectories $(\phi_i(t))$ satisfying $\dot{\phi}(t) \in P$ for all $t$ and $\nu_i(\phi(t)) \leq \dot{\phi}_i(t) \leq \nu_i(\phi(t))$ for each $i$ and all $t$, if there exists a differentiable function $X : P \rightarrow S_n \times S_n$ such that for all $\phi \in P$, the algebraic condition $X(\phi) > 0$, and

$$
\begin{bmatrix}
A^T(\phi)X(\phi) + X(\phi)A(\phi) + \sum_{i=1}^m \nu_i \frac{\partial X}{\partial \phi_i} X(\phi)B(\phi) & C^T(\phi) \\
B^T(\phi)X(\phi) & -I & D^T(\phi) \\
C(\phi) & D(\phi) & -I
\end{bmatrix} < 0 \quad (D.7)
$$

are satisfied for all $\phi \in P$.

Proof: For the clarity of the proof, we will consider systems with only one parameter, i.e., $P \subset \mathbb{R}^1$. Assume that $X(\phi) : P \rightarrow S_n \times S_n$ is differentiable and below three conditions are satisfied for all $\phi \in P$;

$$
X(\phi) > 0 \quad (D.8)
$$

$$
\begin{bmatrix}
A^T(\phi)X(\phi) + X(\phi)A(\phi) + \nu(\phi) \frac{\partial X}{\partial \phi} X(\phi)B(\phi) & C^T(\phi) \\
B^T(\phi)X(\phi) & -I & D^T(\phi) \\
C(\phi) & D(\phi) & -I
\end{bmatrix} < 0 \quad (D.9)
$$

$$
\begin{bmatrix}
A^T(\phi)X(\phi) + X(\phi)A(\phi) + \nu(\phi) \frac{\partial X}{\partial \phi} X(\phi)B(\phi) & C^T(\phi) \\
B^T(\phi)X(\phi) & -I & D^T(\phi) \\
C(\phi) & D(\phi) & -I
\end{bmatrix} < 0 \quad (D.10)
$$

Taking an integer $0 \leq \lambda \leq 1$, we can multiply the Eqn.D.9 by $(1 - \lambda)$ and the Eqn.D.10 by $\lambda$. When we add the resulting matrices

$$
\begin{bmatrix}
A^T(\phi)X(\phi) + X(\phi)A(\phi) + \lambda(\nu(\phi) - \nu_i(\phi)) \frac{\partial X}{\partial \phi_i} X(\phi)B(\phi) & C^T(\phi) \\
B^T(\phi)X(\phi) & -I & D^T(\phi) \\
C(\phi) & D(\phi) & -I
\end{bmatrix} < 0 \quad (D.11)
$$

can be obtained. Note that the freedom in $\lambda$ implies that for any $\kappa(\phi)$ with
\[ \nu(\phi) \leq \kappa(\phi) \leq \overline{\nu}(\phi), \]

\[
\begin{bmatrix}
A^T(\phi)X(\phi) + X(\phi)A(\phi) + \kappa(\phi) \frac{\partial X}{\partial \phi} X(\phi)B(\phi) & C^T(\phi) \\
B^T(\phi)X(\phi) & -I & D^T(\phi) \\
C(\phi) & D(\phi) & -I
\end{bmatrix} < 0 \quad (D.12)
\]

inequality must be satisfied for all \( \phi \in \mathcal{P} \). For the parameter trajectory \( \phi(\cdot) \) satisfying \( \phi(t) \in \mathcal{P} \) and \( \nu(\phi(t)) \leq \dot{\phi}(t) \leq \overline{\nu}(\phi(t)) \) for all \( t \), consider the system

\[
\begin{bmatrix}
\dot{x}(t) \\
z(t)
\end{bmatrix} = \begin{bmatrix}
A(\phi(t)) & B(\phi(t)) \\
C(\phi(t)) & D(\phi(t))
\end{bmatrix} \begin{bmatrix}
x(t) \\
\omega(t)
\end{bmatrix} \quad (D.13)
\]

to be evolving. The algebraic conditions in Eqns.D.8, D.9, D.10 holds with \( \phi = \phi(t) \) and \( \kappa = \dot{\phi}(t) \). Hence, for each \( t \),

\[
\begin{bmatrix}
x^T(t) \omega^T(t) z^T(t)
\end{bmatrix} \begin{bmatrix}
X(\phi)A(\phi) + \frac{1}{2} \dot{\phi}(t) \frac{\partial X}{\partial \phi} X(\phi)B(\phi) & 0 \\
0 & -\frac{1}{2} I & 0 \\
C(\phi) & D(\phi) & -\frac{1}{2} I
\end{bmatrix} \begin{bmatrix}
x(t) \\
\omega(t) \\
z(t)
\end{bmatrix} \leq 0 \quad (D.14)
\]

holds. This is just equal to

\[
x^T(t)X(\phi(t))\dot{x}(t) + \frac{1}{2} x^T(t)\dot{\phi}(t) \frac{dX}{d\phi} |_{\phi(t)} x(t) - \frac{1}{2} ||\omega||^2 + \frac{1}{2} ||z||^2 \leq 0 \quad (D.15)
\]

which could be rewritten\(^3\) as

\[
\frac{1}{2} \frac{d}{dt} \left[ x^T(t)X(\phi)x(t) \right] - \frac{1}{2} ||\omega||^2 + \frac{1}{2} ||z||^2 \leq 0 \quad (D.16)
\]

which can be integrated from from \( t = 0 \) to \( t = T \) and assuming \( x(0) = 0 \) to obtain

\[
||z||_{2,T}^2 \leq ||\omega||_{2,T}^2 - ||X^\frac{1}{2}(\phi(t))x(T)||^2 \leq ||\omega||_{2,T}^2 \quad (D.17)
\]

\(^3\)Noting that \( \frac{1}{2} \frac{d}{dt} \left[ x^T(t)X(\phi)x(t) \right] = \frac{1}{2} \left( \dot{x}^T(t)X(\phi)x(t) + x^T(t)\dot{\phi}(t)X(\phi) |_{\phi(t)} x(t) + x^T(t)X(\phi)\dot{x}(t) \right) \)

and the equality of the first and the third terms.
Consequently, for the $L_2$-induced norm

$$
||G_\phi||_{2 \rightarrow 2} = \sup_{||\omega||_2 \leq 1} \frac{||G_\phi \omega||_2}{||\omega||_2} = \sup_{||\omega||_2 \leq 1} \frac{||z||_2}{||\omega||_2} < 1
$$

(D.18)
can be concluded by interpreting the Eqn.D.17. Following the above steps, the proof can be generalized when there is more than one parameter ($m > 1$), $\phi \in \mathbb{R}^m$.

### D.3 Parameter Dependent Optimization

Parameter-dependent optimization is a derivative form of the constant matrix optimization which is the source of many control design algorithms.

**Problem D.3.1 (Parameter-dependent Optimization).**

$$
\min_{Q(\phi) \in \mathbb{R}^{m \times p}} \quad \sigma [R(\phi) + U(\phi)Q(\phi)V(\phi)] < 1
$$

(D.19)

where $R(\phi) \in \mathbb{R}^{l \times l}$, $U(\phi) \in \mathbb{R}^{l \times m}$, $V(\phi) \in \mathbb{R}^{p \times l}$ and $m, p \leq l$.

The necessary and sufficient conditions for the solution of the parameter-dependent optimization given in Problem D.3.1 can be obtained[54] by applying the result of the constant matrix optimization at each point of the parameter set $\phi \in \mathcal{P}$;

Consider $U_\perp \in \mathbb{R}^{l \times (l-m)}$ and $V_\perp \in \mathbb{R}^{(l-p) \times l}$ such that both $[U \ U_\perp]$ and $[V \ V_\perp]$ are invertible and $U^*U_\perp = 0_{m \times (l-m)}$, $VV_\perp = 0_{p \times (l-p)}$ hold. Then,

$$
\inf_{Q(\phi) \in \mathbb{R}^{m \times p}} \quad \sigma[R(\phi) + U(\phi)Q(\phi)V(\phi)] < 1
$$

(D.20)

if and only if

$$
V_\perp(\phi)[R^*(\phi)R(\phi) - I]V_\perp(\phi) < 0
$$

and

$$
U_\perp^*(\phi)[R(\phi)R^*(\phi) - I]U_\perp(\phi) < 0
$$

(D.21)

for all possible parameter set in $\phi \in \mathcal{P}$.

The $Q(\phi)$ matrix can be taken into consideration with respect to Linear Fractional Transformation equation form.
Vita

Muhittin Yılmaz

Muhittin Yılmaz was born in Akdoğan, Kızılahamam, Ankara, Turkey, in January 1971.

He received the B.Sc. degree in Electrical and Electronics Engineering Department at Gazi University, Ankara, Turkey, in July 1992.

When he won a scholarship while pursuing Masters degree, he left for the Penn State University in 1994 and obtained his M.Sc. degree from Pennsylvania State University in 1996, doing research on fatigue mitigation system by using artificial intelligence, under the supervision of Professor Mario Sznaier.

He is now in the Ph.D. program in the Electrical Engineering department of Penn State University, doing research on robust systems theory applications on macroeconomic stabilization problems under the supervision of Professor Mario Sznaier.